

Rapid progress of DNA replication studies in Archaea, the third domain of life

ISHINO Yoshizumi* & ISHINO Sonoko

*Department of Bioscience and Biotechnology, Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University,
Fukuoka 812-8581, Japan*

Received March 27, 2012; accepted April 20, 2012

Archaea, the third domain of life, are interesting organisms to study from the aspects of molecular and evolutionary biology. Archaeal cells have a unicellular ultrastructure without a nucleus, resembling bacterial cells, but the proteins involved in genetic information processing pathways, including DNA replication, transcription, and translation, share strong similarities with those of Eukaryota. Therefore, archaea provide useful model systems to understand the more complex mechanisms of genetic information processing in eukaryotic cells. Moreover, the hyperthermophilic archaea provide very stable proteins, which are especially useful for the isolation of replisomal multicomplexes, to analyze their structures and functions. This review focuses on the history, current status, and future directions of archaeal DNA replication studies.

Archaea, DNA replication, DNA polymerase, replication fork, replisome, hyperthermophile

Citation: Ishino Y, Ishino S. Rapid progress of DNA replication studies in Archaea, the third domain of life. *Sci China Life Sci*, 2012, 55: 386–403, doi: 10.1007/s11427-012-4324-9

The accurate duplication and transmission of genetic information is essential and crucially important for living organisms. The molecular mechanism of DNA replication has been one of the central themes of molecular biology since the discovery of the double helix DNA structure in 1953 [1]. Since then, molecular biologists have been making continuous efforts to elucidate the precise mechanism of DNA replication. We now know the protein factors involved and each of their functions in the basic mechanism; from origin of replication recognition to replication fork progression in a semi-conservative replication manner, using continuous and discontinuous strand syntheses. Archaeal homologs of many eukaryotic replication proteins (including ORC, Cdc6, GINS, MCM, RPA, primase, DNA polymerase, PCNA, RFC, FEN1, and DNA ligase) have been identified (Figure 1) and biochemically characterized [2–4]. Their similarities indicate that the DNA replication machineries of Archaea and

Eukaryota evolved from a common ancestor, which was different from that of Bacteria [5]. In addition, based on predictions from the genome size and the number of genes, the archaeal replication machinery is a simplified form of that in eukaryotes, and is a good model to elucidate the functions of each component of the complex. On the other hand, the circular genome structure is conserved in Bacteria and Archaea. These features have encouraged studies of DNA replication in Archaea, in the hopes of gaining fundamental insights into this process and its machinery from an evolutionary perspective. Active research on archaeal DNA replication started in about 1990, as compared with bacterial studies before 1960 and eukaryotic studies before 1980. However, the progress in archaeal research has been rapid, especially with the available total genome sequences, and now the depth of knowledge of archaeal DNA replication has almost caught up with that of the bacterial and eukaryotic research fields. In this review, we will summarize the research on the archaeal proteins involved in the DNA

*Corresponding author (email: ishino@agr.kyushu-u.ac.jp)

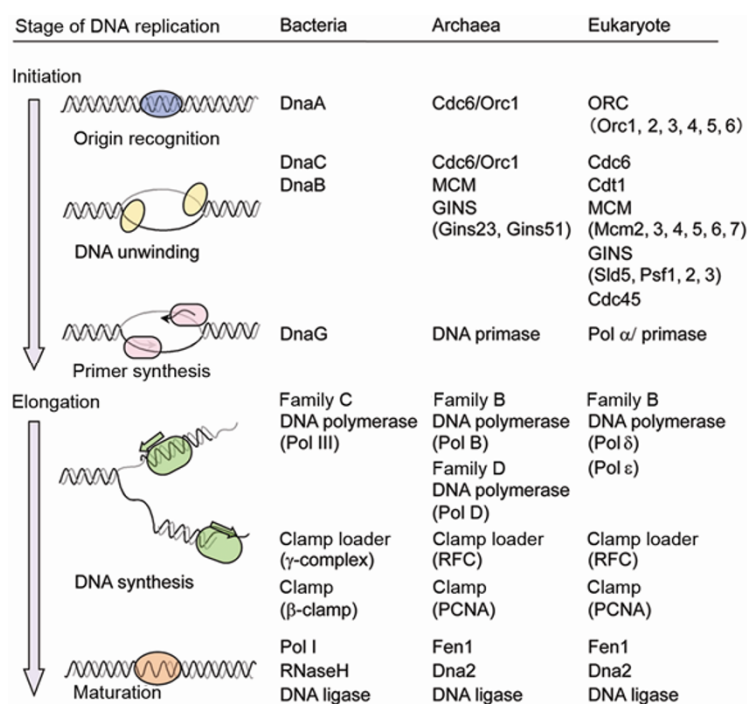


Figure 1 DNA replication proteins and their functions in the three domains of life. Most of the archaeal proteins are designated as homologs of the proteins identified from eukaryotic DNA replication studies.

replication process.

1 DNA replication-related proteins

1.1 Origin of replication (*oriC*) and origin recognition proteins

We identified the first archaeal gene encoding an amino acid sequence similar to those of eukaryotic Cdc6 and Orc1 adjacent to the gene encoding DNA polymerase II (now called PolD, as described below) in the *Pyrococcus furiosus* genome [6]. Subsequently, we identified the origin of replication (*oriC*) just upstream of the gene encoding the Cdc6 and Orc1-like sequences in the *Pyrococcus* genome [7]. We named the gene product Cdc6/Orc1, because of its roughly equal homology to regions of eukaryotic Orc1 and Cdc6, after confirming that the protein actually bound to the *oriC* region on the chromosomal DNA [7]. The archaeal genome sequences had conserved 13 bp repeats characteristic of *oriC*, as predicted by bioinformatics [8]. In *Pyrococcus* genomes, two of the repeats were longer than the conserved 13 bp and surrounded a putative DNA unwinding element (DUE) with an AT-rich sequence (Figure 2) [9]. The longer repeated sequence was designated as an Origin Recognition Box (ORB), and was recognized by Cdc6/Orc1 in a *Sulfolobus solfataricus* study [10]. The 13 base repeat was called a miniORB, being a minimal version of an ORB. We found that *P. furiosus* Cdc6/Orc1 binds to the *oriC* region in the genome with extreme specificity, using a whole genome

microarray analysis, and showed that highly purified Cdc6/Orc1 binds to the ORB and miniORB in the *oriC* region of *P. furiosus* *in vitro* [11].

The crystal structure of the Cdc6/Orc1 protein from *Pyrobaculum aerophilum* [12] and one of the two Cdc6/Orc1 proteins, ORC2 (the authors referred to the two homologs in this organism as ORC1 and ORC2), from *Aeropyrum pernix* [13] were determined. These Cdc6/Orc1 proteins consist of three structural domains. Domains I and II fold in a manner similar to the AAA⁺ family proteins. Domain III has a winged helix (WH) fold, which is present in a number of DNA binding proteins. There are four ORBs arranged in pairs on both sides of the DUE in the *oriC* region of *A. pernix*, and ORC1 binds to each ORB as a dimer. A mechanism was proposed in which after all four ORBs have bound ORC1 proteins, a higher-order assembly of the origin initiates unwinding of the DUE, with alterations in both topology and superhelicity [14]. Furthermore, two reports described the crystal structures of *S. solfataricus* Cdc6-1 and Cdc6-3 (two of the three Cdc6/Orc1 proteins in this organism) forming a heterodimer bound to *ori2* DNA (one of the three origins in this organism) [15], and that *A. pernix* ORC1 bound to an origin sequence [16]. These studies revealed that both the N-terminal AAA⁺ ATPase domain (domain I+II) and C-terminal winged-helix (WH) domain (domain III) contribute to origin DNA binding. In addition, the structural information not only defined the polarity of initiator assembly on the origin, but also indicated the induction of substantial distortion into the DNA strands. This

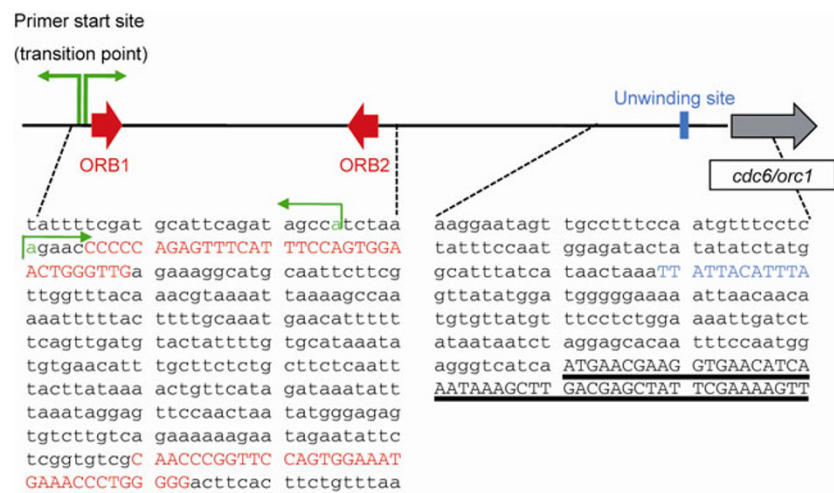


Figure 2 Detailed structure of the *oriC* region in the *Pyrococcus* genome. *oriC*, located in the upstream region of the *cdc6/orc1* gene, is schematically drawn with the exact nucleotide sequences around the first unwinding site, identified by P1 nuclease assays, and the initiation site for primer synthesis, identified by a RIP assay. The consensus sequences of ORB1 and ORB2 (origin recognition box) are indicated in red, and the 12 bp P1-sensitive site is indicated in blue. The region coding for Cdc6/Orc1 is underlined.

Table 1 Distribution of DNA polymerases from seven families in the three domains of life^{a)}

Family	Bacteria		Archaea					Eukaryote	
		Crenarchaeota	Euryarchaeota	Korarchaeota	Aigarchaeota	Thaumarchaeota			
A	Pol I							Pol θ Pol γ ^{**}	
B	Pol II	Pol BI Pol BII	Pol I(B)	Pol BI Pol BII	Pol BII	Pol BII		Pol α, Pol δ Pol ε, Pol ζ	
C	Pol III								
D			Pol D	Pol D	Pol D	Pol D			
E		Pol E [*]							
X								Pol β, Pol λ Pol μ, Pol σ	
Y	Pol IV Pol V							Pol η, Pol ι Pol κ	

a) *, plasmid-coded; **, mitochondrial.

distortion probably triggers the unwinding of the duplex DNA to begin replication. Furthermore, structural data provided the detailed interaction mode between the initiator protein and the *oriC* DNA. Mutational analyses of the *M. thermautotrophicus* Cdc6-1 protein also revealed the essential interaction between an arginine residue conserved in the archaeal Cdc6/Orc1 and an invariant guanine in the ORB sequence [17].

Following the discovery of *P. furiosus* Cdc6/Orc1 in 1995 [6], difficulties in preparing highly purified protein in a soluble fraction made it difficult to characterize the protein. In our recent report using protein prepared in a denaturation-renaturation procedure [18], we found that the *P. furiosus* Cdc6/Orc1 protein alone can unwind duplex DNA at a specific site in the *oriC* region *in vitro*, in an ATP-independent manner. The local unwinding site we identified is approximately 670 bp downstream from the transition site between leading and lagging syntheses (Figure 2), as determined by an *in vivo* replication initiation point (RIP) assay [9]. The replication machinery at the unwound site is still to be established. Although the details of the machinery are

not fully understood in archaea, it should minimally include MCM, GINS, primase, PCNA, DNA polymerase, and RPA, as described below. The assembly of these factors must be within the vicinity of the single-stranded region.

The following studies revealed that the ATPase activity of the Cdc6/Orc1 protein was completely suppressed by binding to DNA containing the origin recognition box (ORB). Limited proteolysis and DNase I-footprint experiments suggested that the Cdc6/Orc1 protein changes its conformation on the ORB sequence in the presence of ATP. The reason for this conformational change is unknown, but it may have an important function in the initiation process [19]. In addition, results from an *in vitro* recruitment assay indicated that Mcm (minichromosome maintenance), the replicative DNA helicase (described below), is recruited into the *oriC* region in a Cdc6/Orc1-dependent, but not ATP-dependent, manner [19]. However, this recruitment is not sufficient for the unwinding function of Mcm, which indicates another protein remains to be identified for the functional loading of this helicase to start the progression of the replication fork DNA.

Table 2 The archaeal DNA ligases characterized to date and their cofactor utilization

Species	Cofactor	Reference
<i>Acidithiobacillus ferrooxidans</i>	ATP	[165]
<i>Aeropyrum pernix</i>	ATP ADP	[166]
<i>Ferroplasma acidarmanus</i>	ATP	[197]
<i>Ferroplasma acidophilum</i>	ATP NAD ⁺	[165]
<i>Methanothermobacterium thermoautotrophicum</i>	ATP	[198]
<i>Picrophilus torridus</i>	ATP NAD ⁺	[165]
<i>Pyrococcus horikoshii</i>	ATP	[169]
<i>Pyrococcus furiosus</i>	ATP	[170]
<i>Staphylothermus marinus</i>	ATP ADP	[167]
<i>Sulfolobococcus zilligii</i>	ATP ADP GTP	[168]
<i>Sulfolobus acidocaldarius</i>	ATP	[165]
<i>Sulfolobus shibatae</i>	ATP	[199]
<i>Thermococcus fumicolans</i>	ATP NAD ⁺	[163]
<i>Thermococcus kodakarensis</i>	ATP NAD ⁺	[162]
<i>Thermococcus sp.</i>	ATP NAD ⁺	[164]
<i>Thermococcus sp. 1519</i>	ATP	[200]
<i>Thermoplasma acidophilum</i>	ATP NAD ⁺	[165]

Analyses of the multiple origins (more than two *oriCs* in a single chromosomal DNA), identified on several archaeal genomes, have primarily been performed in *Sulfolobus* strains [10,20–24]. This is also an interesting subject in the research field of archaeal DNA replication. Questions remain regarding how the initiations of replication from multiple origins are regulated, and how the replication forks progress after the collision of two forks from opposite directions.

1.2 MCM helicase

Another important protein involved in the initiation of DNA replication is the helicase, which unwinds duplex DNA for replication fork progression. The MCM protein complex, consisting of six subunits (Mcm2–7), is the replicative helicase “core” in eukaryotic cells [25]. The MCM complex further forms a higher complex with Cdc45 and GINS (Sld5-Psf1-Psf2-Psf3), and this complex, called the CMG complex, is now believed to be the functional helicase in eukaryotic cells (Figure 3) [26].

Most archaeal genomes appear to encode at least one Mcm homolog, and the helicase activities of these proteins from several archaeal organisms have been confirmed *in vitro* [27–31]. In contrast to the eukaryotic MCM complex, these archaeal MCMs, consisting of a hexamer or double hexamer, have distinct DNA helicase activity by themselves *in vitro*. Therefore, these MCMs may function on their own as the replicative helicase *in vivo*. The structure-function relationships of the archaeal Mcms have been aggressively studied [32]. We reported that Mcm preferentially binds to the origin *in vivo*, as determined using a chromatin im-

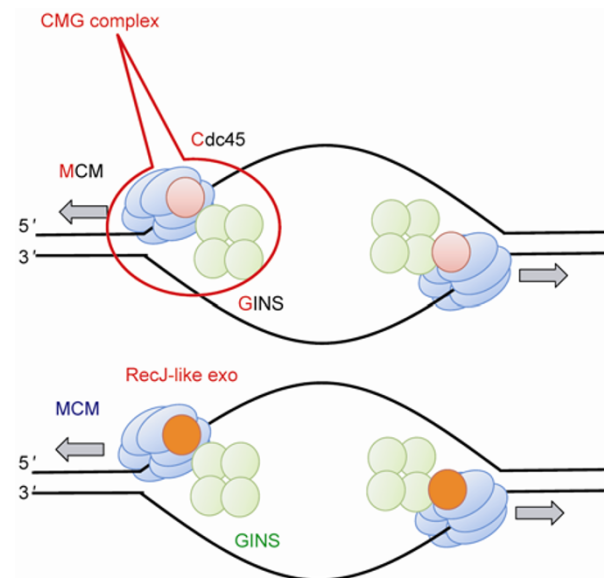


Figure 3 Replicative helicase complexes in eukaryotes and archaea. The CMG complex is the core of the helicase for the template DNA unwinding reaction in eukaryotes. The archaeal genomes contain the homologs of the Mcm and GINS proteins, but a Cdc45 homolog has not been identified. Recent research suggests that a RecJ-like exonuclease, with weak sequence homology to that of Cdc45, may work as a helicase complex with MCM and GINS, although the exact role of its 5′–3′ exonuclease remains unknown.

munoprecipitation (ChIP) method in *P. furiosus*, especially in the exponentially growing phase [7,11]. We have also been studying the MCM helicase of *P. furiosus* *in vitro*, and could not detect significant helicase activity from the purified protein. However, the DNA helicase activity was clearly stimulated by the addition of GINS (the Gins23-Gins51 complex), which is the homolog of the eukaryotic GINS complex (described below in more detail). This result suggests that GINS works as an accessory factor to form a replication fork progression complex (RPC) in *P. furiosus* [31].

The interactions between the Cdc6/Orc1 and Mcm proteins have been reported in several archaeal organisms, including *P. furiosus*, as described above. It was found that the two Cdc6/Orc1 homologs, Cdc6-1 and Cdc6-2, both inhibit the helicase activity of MCM (Mcm complex) in *Methanothermobacter thermoautotrophicus* [33,34]. The Cdc6-1 protein also inhibits the helicase activity of MCM in *S. solfataricus* [35]. In contrast, the Cdc6-2 protein stimulates the helicase activity of Mcm in *Thermoplasma acidophilum* [36].

One interesting feature of archaea is that several species have multiple chromosomal genes encoding Mcm homologs. Recent comprehensive genomic analyses revealed that 13 archaeal species have more than one *mcm* gene. Genetic context analysis showed that many of the *mcm* genes in archaeal genomes reside within mobile elements, originating from viruses and plasmids [37]. For example, *Thermococ-*

cus kodakarensis has three *mcm* genes [38]. Two of the three genes are located in regions where genetic elements have presumably been integrated into the genome. However, these genes are apparently stably inherited, and their gene products may perform some important functions in the DNA metabolism of *T. kodakarensis*. The establishment of a genetic manipulation system, the first for a hyperthermophilic euryarchaeon, in *T. kodakarensis* [39,40], is an advantage for investigating the function of the Mcm proteins. Gene disruption experiments were recently performed for each *mcm* gene by two groups, including ours [41,42]. These experiments revealed that *mcm3* cannot be disrupted, in contrast to the other two *mcm* genes which were easily knocked out. This fact, as well as its relative abundance in the cells, indicates that Mcm3 is the main helicase core protein in the normal DNA replication process in *T. kodakarensis*. The stability of the hexameric structure in solution, found only for Mcm3, also supports the proposal that the Mcm3 protein is the predominant helicase in DNA replication.

Therefore, the function of the other two Mcm proteins remains unclear. The research group of Reeve and Kelman recently reported an investigation of the network between DNA replication-related proteins in *T. kodakarensis* [43]. They introduced a His-tag sequence into 19 different genes in the *T. kodakarensis* genome that encode DNA replication-related proteins. Protein complexes were purified from exponentially growing cells of each strain by His-tag affinity chromatography (Ni^{2+} -chelating column), and the fractions were subjected to mass spectrometry (MS) analysis to detect the proteins included in each complex. They detected the complex from a strain expressing the His-tagged Mcm2 protein, indicating that the His-tagged Mcm2 could be effectively produced in the strain. Furthermore, they showed that the recombinant His-tagged Mcm2 has ATPase and helicase activities *in vitro*. However, our western blot analysis could not detect TkoMcm2 in the extract from exponentially growing *T. kodakarensis* cells. Our subsequent reverse-transcriptase PCR (RT-PCR) experiment detected *mcm2* gene transcript in the cells (Ishino *et al.*, unpublished result). Therefore, the *mcm2* gene appears to be expressed under normal growth conditions. It is also possible that the expression of *mcm2* is constitutive, but the Mcm2 protein is degraded rapidly. Further experiments to measure the efficiency of *mcm2* gene transcription by quantitative PCR, as well as to assess the stability of the Mcm2 protein in the cell extract, are needed. Our experiments demonstrated robust DNA unwinding activities *in vitro* from two of the three Mcm proteins. The helicase activity of Mcm1 is strong *in vitro*, and a distinct amount of Mcm1 protein is present in *T. kodakarensis* cells. Moreover, Mcm1 functionally interacts with the GINS complex from *T. kodakarensis* [42]. These facts strongly suggest that Mcm1 does participate in some aspect of DNA transactions, and may be substituted with Mcm3. Our immunoprecipitation experiments showed that

Mcm1 co-precipitated with Mcm3 and GINS, although they did not form a heterohexameric complex [42]. This suggested that Mcm1 is involved in the replisome or repair-some, and shares some function in *T. kodakarensis* cells. Further phenotypic analyses investigating the sensitivities of the $\Delta mcm1$ and $\Delta mcm2$ mutant strains to DNA damage caused by various mutagens, as reported for other DNA repair-related genes in *T. kodakarensis* [44], may help elucidate the functions of these Mcm proteins.

In the case of the *Methanococcus maripaludis* S2 strain, which harbors four *mcm* genes (three of which seem to be derived from phage), a shotgun proteomics study detected peptides originating from three out of the four *mcm* gene products [45]. Furthermore, the four gene products co-expressed in *E. coli* were co-purified in the same fraction [46]. These results suggest that multiple Mcm proteins are functional in archaeal cells.

A further important issue is how Mcm is recruited onto the unwound region of *oriC* and whether Cdc6/Orc1 is the sole recruiter. Additional critical recruiting factors may exist in archaeal cells. The relationship between the ATPase activity of Cdc6/Orc1 and the Mcm loading of the *oriC* region is also a key question to be solved. Further studies are required to understand the ultimate principles of DNA replication initiation.

1.3 GINS

The eukaryotic GINS complex consists of four different proteins, Sld5, Psf1, Psf2, and Psf3 (GINS is an acronym for the Japanese *go-ichi-ni-san*, meaning 5-1-2-3, after these four subunits), originally identified in *Saccharomyces cerevisiae* as essential protein factors for the initiation of DNA replication [47]. The amino acid sequences of the four subunits in the GINS complex share some conservation, suggesting that they are ancestral paralogs [48]. Finely detailed bioinformatics analyses identified the homologs of GINS subunits in archaeal genomes; however, most of the archaea have only one gene encoding this family protein, and more interestingly, the crenarchaea and euryarchaea (the two major subdomains of archaea) characteristically have sequences more similar to Psf2 and Psf3, and Sld5 and Psf1, respectively [31,48]. A GINS homolog, designated Gins23, was detected in *S. solfataricus* in a yeast two-hybrid screen for interaction partners of MCM. Another subunit, designated Gins15, was identified by the MS analysis of an immunoaffinity-purified native GINS complex from a *S. solfataricus* cell extract, using an anti-GINS23 antibody [49]. The *S. solfataricus* GINS complex, composed of Gins23 and Gins15, forms a tetrameric structure with a 2:2 molar ratio [49]. We also reported the GINS complex from *P. furiosus*, a complex of Gins23 and Gins51 with a 2:2 ratio, as the first euryarchaeal GINS (we prefer to call the GINS subunit proteins Gins23 and Gins51 (instead of Gins15), consistent with the original nomenclature of GINS from 5-1-2-3 [31].

The Mcm2–7 hexamer was co-purified in complex with Cdc45 and GINS from *Drosophila melanogaster* embryo extracts and *S. cerevisiae* lysates. The Cdc45-MCM2–7-GINS (CMG) complex (Figure 3), as described above, was also associated with the replication fork in *Xenopus laevis* egg extracts, and a large molecular machine, containing Cdc45, GINS, and MCM2–7, was proposed as the unwindosome to separate the DNA strands at the replication fork [50]. These results indicate that GINS is a critical factor for not only the initiation process, but also the elongation process in eukaryotic DNA replication. *S. solfataricus* GINS interacts with MCM and primase, although no stimulation or inhibition of either the helicase or primase activity was observed by the interaction with GINS *in vitro* [49]. Conversely, the DNA helicase activity of *P. furiosus* MCM is clearly stimulated by the addition of the *P. furiosus* GINS complex, as described above [31].

T. acidophilum, like many euryarchaea, has only one gene sequence homologous to Gins51 in its genome. Gel filtration and electron microscopy analyses revealed that *T. acidophilum* Gins51 forms a homotetramer. A physical interaction between *T. acidophilum* Gins51 and Mcm was detected by surface plasmon resonance analysis (SPR). TaGins51 did not affect the helicase activity of its cognate MCM when an equal ratio of each molecule was tested *in vitro* [51]. These results suggest that another factor is required to form a stable helicase complex with MCM and GINS at the replication fork in *T. acidophilum*. However, an excess amount of TaGINS clearly stimulated the helicase activity (Ogino *et al.*, unpublished result). In the case of *T. kodakarensis*, the ATPase and helicase activities of MCM1 and MCM3 were clearly stimulated by *T. kodakarensis* GINS *in vitro*. It is noteworthy that the helicase activity of MCM1 was stimulated more than that of MCM3. Physical interactions between the *T. kodakarensis* Gins and Mcm proteins were also detected [42]. These reports suggested that the MCM-GINS complex is a common part of the replicative helicase in Archaea (Figure 3).

Recently, we determined the crystal structure of the *T. kodakarensis* GINS tetramer, comprising Gins51 and Gins23, and compared it with the reported human GINS structures [52]. Each subunit of human GINS consists of an α -helical main domain and a β -stranded small domain, and assembles into the heterotetramer, which exhibits a unique trapezoidal shape [53–55]. The four subunits share a similar fold, and they are classified into two groups: Sld5 and Psf1 possess the α -helical (A) domain at the N-terminus and the β -stranded domain (B) at the C-terminus (AB-type), while Psf2 and Psf3 are the permuted version (BA-type). The backbone structure of each subunit and the tetrameric assembly of *T. kodakarensis* GINS are similar to those of human GINS. However, the location of the C-terminal B domain of Gins51 is remarkably different between the two GINS structures [52]. Based on crystal structures, a homology model of the homotetrameric GINS from *T. acidophi-*

lum was constructed. A long disordered region inserted between the A and B domains allows the differential positioning of the C-terminal domains and, as a consequence, exclusively leads to the formation of an asymmetric homotetramer, rather than a symmetrical assembly [52].

The Cdc45 protein is ubiquitously distributed from yeast to humans, however its exact role in the CMG complex has not been elucidated. No archaeal homolog of Cdc45 has been identified. A detailed analysis of the primary structure of Cdc45 revealed that eukaryotic Cdc45 and prokaryotic RecJ share a common ancestry [56]. A homolog of the DNA binding domain of RecJ was co-purified with GINS from *S. solfataricus* [49]. We detected stimulation of the 5'–3' exonuclease activity of the RecJ homologs from *P. furiosus* and *T. kodakarensis* by the cognate GINSs (Ishino *et al.*, unpublished result). A recent report found that the RecJ homolog from *T. kodakarensis* forms a stable complex with the *T. kodakarensis* GINS, and that the 5'–3' exonuclease activity is enhanced *in vitro* (therefore, the RecJ homolog was designated GAN, from GINS-associated nuclease) [57]. Another related report found that the human Cdc45 protein could bind single-stranded, but not double-stranded, DNA. The structure obtained by small angle X-ray scattering analysis was consistent with a model compatible with the crystallographic structure of the RecJ family members [58]. These interesting findings will promote further research on the structures and functions of the higher-order unwindosome in archaeal cells (Figure 3).

1.4 Primase

To initiate DNA strand synthesis, a primase is required for the synthesis of a short oligonucleotide primer. DnaG and p48-p58 are well-known primases in Bacteria and Eukaryota, respectively. The p48-p58 primase is further complexed with p180 and p70, in what is called the DNA polymerase α -primase complex. The catalytic subunits of the eukaryotic and archaeal primases, which synthesize a short primer on the template DNA, share some sequence homology with those of the family X DNA polymerases [59]. The first archaeal primase was identified in *Methanococcus jannaschii*, as an ORF with a sequence similar to that of the catalytic subunit, p48, of the eukaryotic primase. The gene product exhibited the ability to synthesize oligonucleotides on the template DNA [60]. We characterized the p48-like protein (p41) from *P. furiosus*. Unexpectedly, the archaeal p41 protein did not catalyze the synthesis of short RNA by itself, but preferentially used deoxynucleotides to synthesize DNA strands up to several kilobases in length [61]. We found that the gene neighboring the p41 gene encodes a protein with some similarity to the p58 subunit of eukaryotic primase. This gene product, named p46, forms a stable complex with p41, and the complex can synthesize a short RNA primer *in vitro* [62], consistent with the fact that the short RNA pri-

mer was identified in *Pyrococcus* cells [9]. However, this primase complex also synthesizes DNA strands several hundred nucleotide in length *in vitro* [62].

Further reports characterizing the homologs from *S. solfataricus* [63–65], *P. horikoshii* [66–68], and *P. abyssi* [69] revealed that these primases have similar properties *in vitro*. Notably, the small subunit, p41 (PriS), is the catalytic subunit, while the large subunit, p46 (PriL), modulates the activity in the heterodimeric archaeal primases. The crystal structure of the N-terminal domain of PriL complexed with PriS of *S. solfataricus* primase revealed that PriL does not directly contact the active site of PriS, and therefore, the large subunit may interact with the synthesized primer to adjust its length to 7–14 mer. The structure of the catalytic center, containing a triple aspartic acid motif, is similar to those of the family X DNA polymerases, as predicted from the amino acid sequence similarities. The 3'-terminal nucleotidyl transferase activity, detected in the *S. solfataricus* primase [63,65], and the gap-filling and strand-displacement activities of the *P. abyssi* primase [69] also support the structural similarity between PriS and the family X enzymes.

A further unique activity, named template-dependent polymerization across discontinuous template (PADT), has recently been reported for *S. solfataricus* PriSL [70]. Huang and co-authors propose that the activity may be involved in double-stranded break repair in Archaea.

Archaeal genomes also encode genes resembling bacterial DnaG primase. We expressed the gene encoding the DnaG homolog from *P. furiosus*, however the protein did not show any primer synthesis activity *in vitro*, and thus the archaeal DnaG-like protein does not appear to be a primase (Fujikane *et al.* unpublished result). The DnaG-like protein was shown to participate in RNA degradation as an exosome component [71,72]. However, a recent report revealed that a DnaG homolog from *S. solfataricus* synthesizes primers 13 nucleotides in length, which is the first evidence for the function of the DnaG homolog in Archaea [73]. It would be interesting to investigate if the two primases, PriSL and DnaG, share the primer synthesis functions for leading and lagging strand replication in archaeal cells, as was suggested by the authors [73]. An interesting hypothesis for the evolution of PriSL and DnaG from the last universal common ancestor (LUCA) has also been proposed by Hu *et al.* [70].

The *Sulfolobus* PriSL protein was shown to interact with MCM through Gins23 [49]. This interaction probably ensures the coupling of DNA unwinding and priming during the replication fork progression [49]. Furthermore, the direct interaction between PriSL and the clamp loader RFC (described below) in *S. solfataricus* may regulate the primer synthesis and its transfer to DNA polymerase in archaeal cells [74].

1.5 Single-stranded DNA binding protein

The single-stranded DNA binding protein, called SSB in Bacteria and RPA (replication protein A) in Archaea and Eukaryota, is important for protecting the unwound single-stranded DNA from nuclease attack, chemical modification, and other disruptions during the DNA replication process. Although there is little amino acid sequence similarity between SSB and RPA, the three-dimensional structures revealed that these proteins have a structurally similar domain containing a common fold, called the oligonucleotide/oligosaccharide-binding (OB) fold [75]. The common structure in SSB and RPA suggests that the mechanism of single-stranded nucleic acid binding is conserved in living organisms. In Bacteria, *E. coli* SSB is a homotetramer of a 20 kD peptide with one OB-fold, and the SSBs from *Deinococcus radiodurans* and *Thermus aquaticus* consist of a homodimer of the peptide containing two OB-folds. In contrast, the eukaryotic RPA is a stable heterotrimer, composed of 70, 32, and 14 kD proteins. RPA70 contains two tandem repeats of an OB-fold, which are responsible for the major interaction with a single-stranded DNA in its central region. The N-terminal and C-terminal regions of RPA70 mediate interactions with many cellular or viral proteins, in addition to RPA32 [76,77]. The middle subunit, RPA32, contains an OB-fold in the central region [78–80]. The C-terminal region of RPA32 interacts with other RPA subunits and various cellular proteins [76,77,81,82]. The smallest subunit, RPA14, also contains an OB-fold [78]. The eukaryotic RPA function has been well studied in the *in vitro* SV40 DNA replication system. RPA interacts with the SV40 T-antigen and the DNA polymerase α -primase complex, forming an initiation complex at the origin of replication [83]. RPA also stimulates Pol α -primase activity and PCNA-dependent Pol δ activity [84,85].

The RPAs from *M. jannaschii* and *M. thermoautotrophicus* were reported in 1998, as the first archaeal single-stranded DNA binding proteins [86–88]. These proteins share amino acid sequence similarity with eukaryotic RPA70, and contain four or five repeated domains with the putative OB-fold and one putative zinc-finger motif. *M. jannaschii* RPA exists as a monomer in solution, and has single-strand DNA binding activity. We characterized the RPA from *P. furiosus* [89]. Similar to the eukaryotic RPA, *P. furiosus* RPA forms a complex consisting of three distinct subunits. This RPA highly stimulated the RadA-promoted strand-exchange reaction *in vitro* [89].

While the euryarchaeal organisms have a eukaryotic-type RPA homolog, the crenarchaeal SSB proteins appear to be much more similar to the bacterial proteins, with a single OB fold and a flexible C-terminal tail. The crystal structure of the SSB protein from *S. solfataricus* was determined to 1.26 Å [90]. The structure of the OB-fold domain is more similar to that of the eukaryotic RPAs, although it was expected to resemble the *E. coli* SSB, supporting the close

relationship between archaea and eukaryotes [90].

The three subunits, RPA1, RPA2, and RPA3, from *Methanosarcina acetivorans* were also characterized [91]. Unlike the multiple RPA proteins found in other archaea and eukaryotes, each of the *M. acetivorans* RPAs can act as a distinct single-stranded DNA-binding protein. The architecture of the *M. acetivorans* RPA also differs from that found in other archaea. Each of the three RPA proteins, as well as their combinations, clearly stimulates the primer extension activity of *M. acetivorans* DNA polymerase BI *in vitro*, as shown previously for bacterial SSB and eukaryotic RPA [91]. Structural analyses of the architectures of SSB and RPA suggested that they are composed of different combinations of the OB-fold. Bacterial and eukaryotic organisms contain one type of SSB or RPA, respectively. In contrast, archaeal organisms have various RPAs, composed of different organizations of OB-folds. *M. acetivorans* RPA1, 2, and 3 have four, two, and two OB-folds, respectively. The hypothesis that homologous recombination might play an important role in generating the diversity of OB-folds in archaea was proposed, based on experiments characterizing the engineered RPAs with various OB-folds [92].

1.6 DNA polymerase

DNA polymerase catalyzes phosphodiester bond formation between the terminal 3'-OH of the primer and the α -phosphate of the incoming triphosphate. DNA polymerases have been classified by amino acid sequence similarity, and seven families, A, B, C, D, E, X, and Y, are now widely recognized (Table 1) [93–97]. The fundamental ability of DNA polymerases to synthesize a deoxyribonucleotide chain is conserved in relation to the structural conservation. However, the more specific properties, including processivity, synthesis accuracy, and substrate nucleotide selectivity, differ among the enzymes. Enzymes within the same family have similar basic properties. *E. coli* has five DNA polymerases, and PolI, PolII, and PolIII belong to families A, B, and C, respectively. PolIV and PolV are classified in family Y, as the DNA polymerases for trans-lesion synthesis (TLS). In eukaryotes, the replicative DNA polymerases, Pol α , Pol δ , and Pol ϵ belong to family B, and the trans-lesion DNA polymerases, η , ι , and κ belong to family Y [98].

Members of the Crenarchaeota have at least two family B DNA polymerases [99,100]. However, there is only one family B DNA polymerase in the Euryarchaeota. Instead, the euryarchaeotic genomes encode a family D DNA polymerase, as described below, which seems specific to these archaeal organisms, as it has never been found in any other living organisms [94,101]. Several archaea have family Y-like DNA polymerases, although they have not been found in many other archaeal organisms. We will not describe the trans-lesion synthesis by family Y DNA polymerase in this review. The role of each DNA polymerase in

archaeal cells is still not known, although the distribution of the DNA polymerases is better understood (Table 1) [102].

We cloned the gene encoding a family B DNA polymerase from *P. furiosus*, the first known entire *pol* gene in archaea [103], while the family B DNA polymerases from *Pyrodicticum occultum* [99] and *A. pernix* [100] were cloned in the 1990s. The most interesting features discovered at the inception of this research area were that archaea possess the eukaryotic Pol α -like (Family B) DNA polymerases [103], and some archaea (crenarchaea) have multiple Pol α -like DNA polymerases, as in eukaryotic cells [99,100]. Lastly, a novel DNA polymerase (Family D) without any sequence homology was also discovered [6]. The early stages of the research on the thermostable DNA polymerases have been summarized in a previous review article [104].

The first family D DNA polymerase was identified from *P. furiosus*, by screening for DNA polymerase activity in the cell extract [105]. Three independent fractions showing deoxynucleotide incorporation activity were obtained from anion-exchange chromatography. DNA polymerase activity, which was not representative of a family B-like enzyme because of its resistance to aphidicolin (α -like DNA polymerase-specific DNA polymerase), was found in one fraction. The corresponding gene was cloned, revealing that this new DNA polymerase consisted of two proteins, named DP1 and DP2, and that the deduced amino acid sequences of these proteins were not conserved in the DNA polymerase families [6]. This DNA polymerase, proposed as PolD from the new family (family D), has only been identified in the euryarchaeal organisms to date [94,101]. *P. furiosus* PolD exhibits efficient strand extension activity and strong proofreading activity [6,106]. Other family D DNA polymerases were also characterized by several groups [107–113].

A genetic study of *Halobacterium* sp. NRC-1 showed that both PolB and PolD are essential for viability [114]. An interesting issue is to elucidate whether PolB and PolD work together at the replication fork for the synthesis of the leading and lagging strands, respectively. According to its biochemical properties, including the usage of an RNA primer and the presence of strand displacement activity, PolD may catalyze lagging strand synthesis [102,112].

The PolD genes were found only in Euryarchaeota, as described above. However, recent environmental genomics and cultivation efforts revealed novel phyla in Archaea: Thaumarchaeota, Korarchaeota, and Aigarchaeota, and their genome sequences showed that Thaumarchaeota and Aigarchaeota harbor the genes encoding PolD and crenarchaeal PolBII [115,116], while Korarchaeota encodes PolBI, PolBII and PolD [117]. In spite of accumulating information about the distribution of DNA polymerases among novel archaeal phyla (Table 1), only one gene for PolBII from the psychrophilic marine thaumarchaeon, *Cenarchaeum symbiosum*, has been analyzed [118]. Biochemical characterization of these gene products will contribute to

research on the evolution of DNA polymerases in living organisms. A hypothesis that the archaeal ancestor of eukaryotes encoded three DNA polymerases, two distinct family B polymerases and a family D polymerase, which all contributed to the evolution of the eukaryotic replication machinery, consisting of Pol α , Pol δ , and Pol ϵ , has been proposed [119].

One unique feature of the archaeal family B DNA polymerases is their specific interactions with uracil and hypoxanthine, which stall the progression of the polymerases on DNA template strands [120]. This phenomenon, called “read-ahead recognition” of deaminated bases, was originally proposed to be characteristic of hyperthermophilic archaea, to protect them against increased deamination at high temperature. However, the family B DNA polymerase from the mesophilic archaeon, *M. acetivorans*, also recognizes uracil and hypoxanthine as efficiently as the enzymes from hyperthermophiles [121]. The crystal structure of the DNA polymerase revealed that read-ahead recognition occurs by an interaction with the deaminated bases in an N-terminal binding pocket, existing specifically in the archaeal family B DNA polymerases [122].

An interesting protein was identified in the plasmid pRN1, isolated from a *Sulfolobus* strain [123]. This protein, ORF904 (named RepA), has primase and DNA polymerase activities in the N-terminal domain, and helicase activity in the C-terminal domain, and is likely to be essential for the replication of pRN1 [124,125]. The amino acid sequence of the N-terminal domain lacks homology to any known DNA polymerase or primase, although the C-terminal domain contains a sequence conserved in the superfamily 3 helicases, and therefore, family E was proposed as a new DNA polymerase family. Similar proteins are encoded by various archaeal and bacterial plasmids, as well as by some bacterial viruses [126]. Recently, protein t2-12p, encoded in the plasmid pTN2 isolated from *Thermococcus nautilus*, was experimentally identified as a DNA polymerase from this family [127]. This enzyme is likely to be responsible for plasmid replication. Further investigation of this family of DNA polymerases will be interesting from an evolutionary perspective.

1.7 PCNA and RFC

The sliding clamp is the processivity factor that retains the DNA polymerase on the template DNA strand for continuous strand synthesis. The doughnut-shaped ring structure of clamp molecules is conserved among living organisms. The eukaryotic and archaeal proliferating cell nuclear antigens (PCNAs) form a homotrimeric ring structure [128,129] that encircles the DNA strand and anchors the DNA polymerase onto the DNA by binding it on one surface (front side) of the ring, as described below (Figure 4). To load the clamp ring onto the DNA strand, a clamp loader is required to interact with the clamp and open its ring to introduce the

DNA strand into the central hole. The γ -complex in Bacteria and replication factor C (RFC) in Archaea and Eukaryota are clamp loaders (Figure 4). We have been studying the structures and functions of *P. furiosus* PCNA [130–134] and RFC [135–138] and found that these molecules are essential for DNA polymerase to perform processive DNA synthesis. RFC loads PCNA onto the DNA strand in an ATP-dependent manner. The mechanism of clamp loading is an interesting subject, and several research groups are actively investigating it [139] (Figure 4). We detected an intermediate PCNA-RFC-DNA complex, in which the PCNA ring is opened, by single particle analysis of electron microscopic images using *P. furiosus* proteins (Figure 4) [140]. These results contributed to our understanding of the molecular mechanism of the clamp loading process [140]. The crystal structure of the complex, including the ATP-bound clamp loader, the ring-opened clamp, and the template-primer DNA, using proteins from bacteriophage T4, has recently been published [141].

Following clamp loading, DNA polymerase accesses the PCNA, and the polymerase-clamp complex catalyzes processive DNA synthesis. Therefore, structural and functional studies of the DNA polymerase-PCNA interaction are the next target to elucidate the overall mechanisms of replication fork progression. Extensive studies of the PCNA-interacting proteins revealed that the proteins contain a small conserved sequence motif, called the PCNA interact-

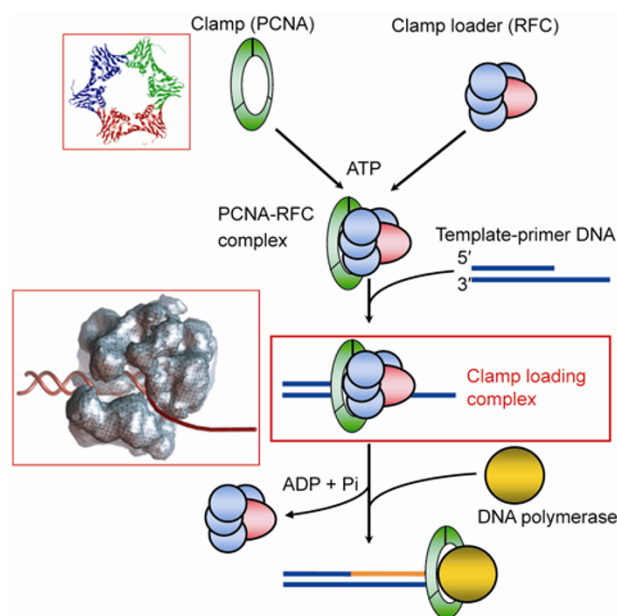


Figure 4 Mechanisms of clamp loading and processive DNA synthesis. The clamp loader (RFC) opens and recruits the clamp (PCNA) ring onto the primer terminus of the DNA strand, in an ATP-dependent manner. The clamp loader is then replaced by DNA polymerase, which can synthesize the DNA strand processively without falling off. The crystal structure of *P. furiosus* PCNA (PDB code 1GE8) and the structure of the clamp loading complex (PCNA-RFC-DNA ternary complex), determined by EM single particle analysis, are shown on the left. Each subunit of the PCNA ring is shown in a different color.

ing protein box (PIP box), which binds to a common site on PCNA [142]. The PIP box consists of the sequence “Qxxhxxaa”, where “x” represents any amino acid, “h” represents a hydrophobic residue (e.g., L, I or M), and “a” represents an aromatic residue (e.g., F, Y or W). Archaeal DNA polymerases have PIP box-like motifs in their sequences [143]. However, only a few studies have experimentally investigated the function of the motifs.

We determined the crystal structure of *P. furiosus* PolB complexed with a monomeric PCNA mutant, which allowed us to construct a convincing model of the polymerase-PCNA ring interaction [144]. Our structural study revealed that a novel interaction is formed between a stretched loop of PCNA and the thumb domain of PolB, in addition to the authentic PIP box. A comparison of the model structure with the previously reported structures of a family B DNA polymerase from RB69 phage, complexed with DNA [145,146], suggested that the second interaction site plays a crucial role in switching between the polymerase and exonuclease modes, by inducing a PCNA-polymerase complex configuration that favors synthesis over editing. This putative mechanism for the control of fidelity of replicative DNA polymerases is supported by experiments in which mutations at the second interaction site enhanced the exonuclease activity in the presence of PCNA [146]. Furthermore, we presented the three-dimensional structure of the DNA polymerase-PCNA-DNA ternary complex, solved by electron microscopic (EM) single particle analysis [147]. This structural view revealed the entire domain configuration of the trimeric ring of PCNA and DNA polymerase, including the protein-protein or protein-DNA contacts (Figure 5). This architecture provides clearer insights into the

switching mechanism between the editing and synthesis modes [147].

In contrast to almost all organisms in Euryarchaeota, which have a single PCNA homolog forming a homotrimeric ring structure, the majority of Crenarchaeota have multiple PCNA homologs, which are capable of forming heterotrimeric rings for their functions [148,149]. It is especially interesting that the three PCNAs, PCNA1, PCNA2, and PCNA3, specifically bind PCNA binding proteins, including DNA polymerases, DNA ligases, and FEN-1 endonuclease [149,150]. Detailed structural studies of the heterologous PCNA from *S. solfataricus* revealed that the interaction modes between the subunits are conserved with those of the homotrimeric PCNAs [151,152].

T. kodakarensis is the only euryarchaeal species that has two genes encoding PCNA homologs in the genome [38]. We cloned these two genes from the *T. kodakarensis* genome, and the highly purified gene products, PCNA1 and PCNA2, were characterized (Kuba *et al.*, in review). *In vivo* analyses revealed that the gene encoding PCNA2 could be knocked out, whereas no PCNA1 deletion mutants could be obtained, suggesting that PCNA1 is essential for DNA replication, and PCNA2 may play a different role in *T. kodakarensis*. A recent report showed that both PCNA1 and PCNA2 form a stable ring structure and work as a processivity factor for *T. kodakarensis* PolB *in vitro*. The crystal structures of the two PCNAs revealed the different interactions at the subunit-subunit interfaces [153]. It will be very interesting to determine how the two PCNAs function together in *T. kodakarensis* cells.

The eukaryotic RFC is a heteropentameric complex, consisting of five different proteins, RFC1–5, in which RFC1 is larger than the other four RFCs. Conversely, the archaeal RFC consists of two proteins, RFCS (small) and RFCL (large), in a 4:1 ratio. Thus, similar to the eukaryotic RFC, most of the archaeal RFCs form a pentameric complex. However, a different form of RFC, consisting of three subunits, RFCS1, RFCS2, and RFCL, in a 3:1:1 ratio, was originally identified from *M. acetivorans* [154]. This RFC stimulates DNA synthesis by a cognate DNA polymerase in the presence of its PCNA *in vitro* [154]. The three subunits of RFC may represent an intermediate stage in the evolution of the more complex RFC in Eukaryotes from the less complex RFC in Archaea [154,155]. The subunit organization and the spatial distribution of the subunits in the *M. acetivorans* RFC complex were analyzed and compared with those of the *E. coli* γ -complex, which is also a pentamer consisting of three different proteins. These two clamp loaders adopt similar subunit organizations and spatial distributions, but the functions of the individual subunits are likely to be diverse [155]. Detailed evolutionary analyses of the archaeal RFC proteins, including the Mcm and PCNA proteins, were published recently [156].

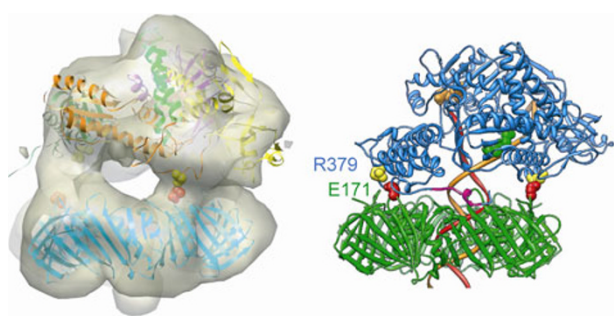


Figure 5 Three-dimensional structure of the PolB-PCNA-DNA complex. The three-dimensional structural model of the PolB-PCNA-DNA ternary complex was obtained by EM single particle analysis, and the crystal structure of each component was docked onto the complex structure. PolB interacts with PCNA primarily at two sites, a conventional PIP box and a newly identified site (around R379). Based on this structure and the related biochemical analyses, a switching mechanism for PolB between the polymerase (synthesizing) and exonuclease (editing) modes was proposed. The second contact between the two molecules (at R379 of PolB and E171 of PCNA), as shown here, appears to be responsible for the configuration of the editing mode-specific complex. A different EM image with a detached second contact, indicating the synthesizing mode, was also observed (data not shown).

1.8 DNA ligase

DNA ligase is essential to connect the Okazaki fragments of discontinuous strand synthesis during DNA replication, and universally exists in all living organisms. This enzyme catalyzes nick-sealing reactions via three nucleotidyl transfer steps, as described in previous review articles [157,158]. In the first step, DNA ligase forms a covalent enzyme-AMP intermediate, by reacting with ATP or NAD⁺ as a cofactor. In the second step, DNA ligase recognizes the substrate DNA, and the AMP is subsequently transferred from the ligase to the 5'-phosphate terminus of the DNA, to form a DNA-adenylate intermediate (AppDNA). In the final step, the 5'-AppDNA is attacked by the adjacent 3'-hydroxy group of the DNA, to form a phosphodiester bond. Generally, DNA ligases are grouped into two families according to their requirement for ATP or NAD⁺ as a nucleotide cofactor in the first reaction step. ATP-dependent DNA ligases are widely found in all three domains of life, whereas NAD⁺-dependent DNA ligases exist mostly in Bacteria. NAD⁺-dependent enzymes are also present in halophilic archaea [159] and some eukaryotic viruses [160].

Three genes (*LIG1*, *LIG3* and *LIG4*) encoding ATP-dependent DNA ligases have been identified in the human genome to date. Human DNA ligase I (*LIG1*), encoded by *LIG1*, is a replicative enzyme that joins Okazaki fragments during DNA replication. The first gene encoding a eukaryotic-like ATP-dependent DNA ligase was found in the thermophilic archaeon *Desulfolobus ambivalens* [161]. Subsequent identification of DNA ligases from archaeal organisms revealed that these enzymes primarily use ATP as a cofactor. However, this classification may not be so strict; the use of NAD⁺, as well as ATP, as a cofactor has been observed in several DNA ligases, including those from *T. kodakarensis* [162], *T. fumicolans*, *P. abyssi* [163], *Thermococcus* sp. NA1 [164], *T. acidophilum*, *Picrophilus torridus*, and *Ferroplasma acidophilum* [165] (Table 2). The dual co-factor specificity (ATP/NAD⁺) is interesting in terms of understanding the structures and functions of these DNA ligases, although ATP is evidently preferable in all of the cases. Further dual co-factor specificity exists in the archaeal DNA ligases, which use both ADP and ATP, as found in the enzymes from *A. pernix* [166] and *Staphylothermus marinus* [167]. In the case of *Sulfobococcus zilligii*, GTP is also the functional cofactor [168]. The DNA ligases from *P. horikoshii* [169] and *P. furiosus* [170] have a strict ATP preference. Sufficient data have not yet been obtained to fully resolve the issue of dual co-factor specificity, so further biochemical and structural analyses are required.

We determined the crystal structure of *P. furiosus* DNA ligase [171] and investigated the physical and functional interactions between the DNA ligase and PCNA [170]. The detailed mode of interaction between human *LIG1* and PCNA is somewhat unclear, because of several conflicting

reports [172–174]. The stimulatory effect of *P. furiosus* PCNA on the enzyme activity of the cognate DNA ligase was observed at a high salt concentration, at which a DNA ligase alone cannot bind to a nicked DNA substrate. Furthermore, we identified the amino acid residues that are critical for the PCNA binding, based on mutational analyses. Interestingly, the binding site is located in the middle of the N-terminal DNA binding domain (DBD), and the binding motif, QKSFF, which we proposed as a shorter version of the PIP box, is actually looped out from the protein surface. We subsequently confirmed that the Gln and the first Phe in the QKSFF motif are especially important for stable binding with PCNA [170]. Interestingly, this motif is located in the middle of the protein chain, rather than at the N- or C-terminal region, where the PIP boxes are usually found. To confirm that this motif is conserved in the archaeal/eukaryotic DNA ligases, we showed that the physical and functional interactions between *A. pernix* DNA ligase and PCNA depend mainly on the phenylalanine residue, which is located in the region predicted by the multiple sequence alignment of the ATP-dependent DNA ligases [175]. It is noteworthy that this shorter version of the PIP box also exists in an archaeal uracil DNA glycosylase (UDG), identified in our continuing work [176].

Human *LIG1*, complexed with DNA, was the first reported crystal structure of an ATP-dependent cellular DNA ligase, although the ligase was an N-terminal truncated form [177]. The enzyme comprises the N-terminal DNA binding domain, the middle adenylation domain, and the C-terminal OB-fold domain. The crystal structure of *LIG1* (residues 233–919), in complex with a nicked, 5'-adenylated DNA intermediate, revealed that the enzyme redirects the path of the double-stranded DNA to expose the nick termini for the strand-joining reaction. Like PCNA, the N-terminal DNA-binding domain works to encircle the DNA substrate and to stabilize it in a distorted structure, positioning the catalytic core on the nick. Our crystal structure of the full length DNA ligase from *P. furiosus* [170] revealed that the architecture of each domain resembles those of *LIG1*, but the domain arrangements differ markedly between the two enzymes. This domain rearrangement is likely derived from the “domain-connecting” role of the helical extension, conserved at the C-termini of the archaeal and eukaryotic DNA ligases. The DNA substrate in the open form of *LIG1* is replaced by motif VI at the C-terminus in the closed form of *P. furiosus* ligase. Both the shapes and electrostatic distributions are similar between motif VI and the DNA substrate, suggesting that motif VI in the closed state mimics the incoming substrate DNA. Subsequently, another crystal structure of the ATP-dependent DNA ligase from *S. solfataricus* was published [178]. This crystal structure is the fully open DNA ligase structure, in which the three domains are highly extended. In this work, the *S. solfataricus* ligase-PCNA complex was also analyzed by small-angle X-ray scattering. *S. solfataricus* ligase bound to the PCNA ring

retains an open, extended conformation. The closed, ring-shaped conformation observed in the LIGI structure is probably the active form to catalyze a DNA end-joining reaction. Therefore, it was proposed that the open-to-closed conformational change occurs for ligation, and the switch in the conformational change is accommodated by a malleable interface with PCNA, which serves as an efficient platform for DNA ligation [178], as illustrated in Figure 6.

Following the publication of these crystal structures, we presented the three-dimensional structure of the ternary complex, consisting of DNA ligase-PCNA-DNA, obtained by EM single particle analysis (Figure 6) [179]. In the complex structure, the three domains of the crescent-shaped *P. furiosus* ligase surround the central DNA duplex, encircled by the closed PCNA ring. The relative orientations of the ligase domains, which differ remarkably from those of the previous crystal structures, suggested that a large domain rearrangement occurs upon ternary complex formation. The DNA ligase contacted PCNA at two sites, the conventional PIP box and a second novel contact in the middle adenylation domain. It is also interesting that a substantial DNA tilt from the PCNA ring axis was observed. Based on the structural model, we considered a mechanism in which the PCNA binding proteins are bound and released sequentially. In fact, most of the PCNA binding proteins share the same binding sites in the interdomain connecting loop (IDCL) and the C-terminal tail of the PCNA. The structural features exclude the possibility that the three proteins contact the

single PCNA ring simultaneously, because DNA ligase occupies two of the three subunits of the PCNA trimer (Figure 6). In the case of our previous analysis of the RFC-PCNA-DNA complex, RFC entirely covers the PCNA ring, thus blocking the access of other proteins (Figure 4) [140]. The ternary complex appears to favor a mechanism involving the sequential binding and release of replication factors. In solution, a DNA ligase can adopt various conformations between the extended and closed forms, as described above. The ternary structure described by our group could be regarded as an intermediate state, just prior to DNA ligation (Figure 6).

1.9 Flap endonuclease 1

Efficient processing of Okazaki fragments is vital for DNA replication and cell proliferation. The primase-synthesized RNA/DNA primers need to be removed to join the Okazaki fragments into an intact lagging strand DNA. Flap endonuclease 1 (FEN1) is primarily responsible for this task. Okazaki fragment maturation is highly coordinated, and the interactions of DNA polymerase δ , FEN1, and DNA ligase with PCNA allow these enzymes to act sequentially during the maturation process, as described above.

FEN1 is a structure-specific 5'-endonuclease that specifically recognizes a double-stranded DNA with an unannealed 5'-flap [180,181]. In Eukaryota, 5'-flap DNA structures are formed by the strand displacement activity of DNA polymerase δ during Okazaki fragment processing and long patch base excision repair. DNA ligase I seals the nick after the flapped DNA is cleaved by FEN1. These processing steps are facilitated by PCNA [182]. The interactions between eukaryotic FEN1 and PCNA have been well characterized [142,173]. The stimulatory effect of PCNA on the FEN1 activity was also examined [183]. The crystal structure of the human FEN1-PCNA complex revealed three FEN1 molecules bound to one PCNA trimer at the C-terminal PIP box (one FEN1 for each subunit), in different configurations [184].

The eukaryotic homologs of FEN1 were found in Archaea [185]. Many structural analyses of archaeal FEN1 have been reported to date. The crystal structures of FEN1 from *M. jannaschii* [186], *P. furiosus* [187], *P. horikoshii* [188], *A. fulgidus* [189], and *S. solfataricus* [151] have been determined. In addition, detailed biochemical studies were performed on *P. horikoshii* FEN1 [190,191]. Thus, studies of the archaeal FEN1 proteins have provided important insights into the structural basis of the cleavage reaction of the flapped DNA. We cloned and purified *P. furiosus* FEN1, and examined its interaction with PCNA in detail. The flap endonuclease activity of *P. furiosus* FEN1 was stimulated by PCNA. Furthermore, the stimulatory effect of PCNA on the sequential action of FEN1 and ligase was observed *in vitro*. Based on these results, a model of the molecular switching mechanism of the last steps of Okazaki-fragment

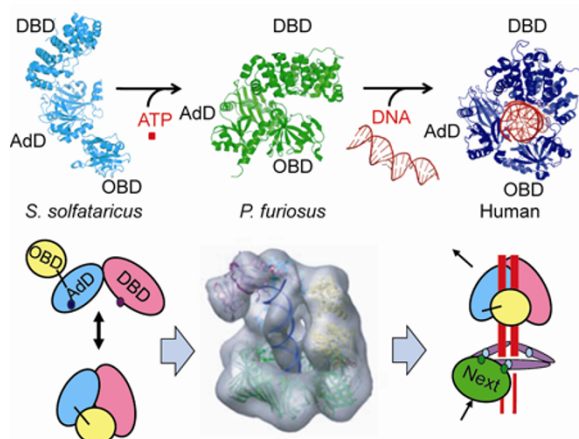


Figure 6 Conformational changes of DNA ligase, according to the ligation process. The ATP-dependent DNA ligases, consisting of three distinct domains, flexibly change their conformations to capture the DNA strand and to process ligation. Three crystal structures supporting these conformational changes are shown. The DNA ligases from *S. solfataricus* (PDB code 2HIV, cyan), *P. furiosus* (PDB code 2CFM, green), and human LIGI with DNA (PDB code 1X9N, dark blue and red) are shown. The predicted ligation process by DNA ligase, with the help of PCNA, is depicted at the bottom. Upon binding to PCNA and DNA, the DNA ligase forms the crescent configuration, which embraces the DNA, as observed from the EM single particle analysis. The DNA ligase then grips the DNA to accomplish the ligation reaction, as observed in the crystal structure of human LIGI-DNA. Accordingly, the binding sites are released, thereby enabling the PCNA ring to interact with the PIP of the next enzyme.

maturation was constructed.

2 Conclusion and perspectives

Research into the molecular mechanism of DNA replication has been a central theme of molecular biology, and studies on archaeal organisms became popular in the whole genome sequencing age. Archaeal studies are especially interesting to understand the mechanisms by which cells live in extreme environmental conditions. It is also noteworthy that the proteins from the hyperthermophilic archaea are more stable than those from mesophilic organisms, and they are especially advantageous for the structural and functional analyses of higher-ordered complexes, such as the replisome. To understand each of the events involved in DNA metabolism, studies on the higher-order complexes, rather than single proteins, are essential. Archaeal research will continue to contribute to the development and advancement of the DNA replication research field, as summarized in a recent review [192].

In addition to basic molecular biology research, DNA replication proteins from thermophiles have been highly useful reagents for gene manipulations, including genetic diagnosis, forensic DNA typing, detection of bacterial and viral infections, as well as basic research. Numerous archaeal enzymes have been developed globally as commercial products, which are utilized daily in various processes. An example of the successful engineering of a DNA polymerase for PCR is the creation of the fusion protein of *P. furiosus* PolB and a nonspecific dsDNA binding protein from *S. solfataricus*, Sso7d, by genetic engineering techniques [193]. To overcome the low processivity of PolB, a small protein with high affinity to the DNA strand was fused to the C-terminus of the enzyme. This artificial DNA polymerase possesses high processivity, while maintaining high fidelity for DNA synthesis. As another example, we successfully developed a novel processive PCR method, using archaeal PolB with the help of a mutant PCNA [194,195]. Improved DNA sequencing technologies, referred to as “next-generation sequencing” have been developed [196,197] and single-molecule detection, using dye-labeled modified nucleotides and longer read lengths, known as “third-generation DNA sequencing” [198] is also available. All of these technologies include DNA polymerases or DNA ligases from various sources including archaeal organisms, indicating that DNA replication enzymes are indispensable for the development of DNA manipulation technology. Further efforts will provide valuable enzymes with specialized activities for promoting genetic engineering technologies, and thermostable enzymes with their associated factors have great potential.

Culture, Sports, Science, and Technology of Japan, the Japan New Energy and Industrial Technology Development Organization, and the Japan Science and Technology Agency. We apologize to the researchers whose work was not cited because of space limitations.

- 1 Watson J D, Crick F H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 1953, 171: 737–738
- 2 Kelman Z, White M F. Archaeal DNA replication and repair. *Curr Opin Microbiol*, 2005, 8: 669–676
- 3 Barry E R, Bell S D. DNA replication in the archaea. *Microbiol Mol Biol Rev*, 2006, 70: 876–887
- 4 Wigley D B. ORC proteins: marking the start. *Curr Opin Struct Biol*, 2009, 19: 72–78
- 5 Leipe D D, Aravind L, Koonin E V. Did DNA replication evolve twice independently? *Nucleic Acids Res*, 1999, 27: 3389–3401
- 6 Uemori T, Sato Y, Kato I, et al. A novel DNA polymerase in the hyperthermophilic archaeon, *Pyrococcus furiosus*: gene cloning, expression, and characterization. *Genes Cells*, 1997, 2: 499–512
- 7 Matsunaga F, Forterre P, Ishino Y, et al. In vivo interactions of archaeal Cdc6/Orc1 and minichromosome maintenance proteins with the replication origin. *Proc Natl Acad Sci USA*, 2001, 98: 11152–11157
- 8 Lopez P, Philippe H, Myllykallio H, et al. Identification of putative chromosomal origins of replication in Archaea. *Mol Microbiol*, 1999, 4: 883–886
- 9 Matsunaga F, Norais C, Forterre P, et al. Identification of short ‘eukaryotic’ Okazaki fragments synthesized from a prokaryotic replication origin. *EMBO Rep*, 2003, 4: 154–158
- 10 Robinson N P, Dionne I, Lundgren M, et al. Identification of two origins of replication in the single chromosome of the archaeon *Sulfolobus solfataricus*. *Cell*, 2004, 116: 25–38
- 11 Matsunaga F, Glatigny A, Mucchielli-Giorgi M H, et al. Genomewide and biochemical analyses of DNA-binding activity of Cdc6/Orc1 and Mcm proteins in *Pyrococcus* sp. *Nucleic Acids Res*, 2007, 35: 3214–3222
- 12 Liu J, Smith C L, DeRyckere D, et al. Structure and function of Cdc6/Cdc18: implications for origin recognition and checkpoint control. *Mol Cell*, 2000, 6: 637–648
- 13 Singleton M R, Morales R, Grainge I, et al. Conformational changes induced by nucleotide binding in Cdc6/ORC from *Aeropyrum pernix*. *J Mol Biol*, 2004, 343: 547–557
- 14 Grainge I, Gaudier M, Schuwirth B S, et al. Biochemical analysis of a DNA replication origin in the archaeon *Aeropyrum pernix*. *J Mol Biol*, 2006, 363: 355–369
- 15 Dueber E L, Corn J E, Bell S D, et al. Replication origin recognition and deformation by a heterodimeric archaeal Orc1 complex. *Science*, 2007, 317: 1210–1213
- 16 Gaudier M, Schuwirth B S, Westcott, S L, et al. Structural basis of DNA replication origin recognition by an ORC protein. *Science*, 2007, 317: 1213–1216
- 17 Majernik A I, Chong I P. A conserved mechanism for replication origin recognition and binding in archaea. *Biochem J*, 2008, 409: 511–518
- 18 Matsunaga F, Takemura K, Akita M, et al. Localized melting of duplex DNA by Cdc6/Orc1 at the DNA replication origin in the hyperthermophilic archaeon *Pyrococcus furiosus*. *Extremophiles*, 2009, 14: 21–31
- 19 Akita M, Adachi A, Takemura K, et al. Cdc6/Orc1 from *Pyrococcus furiosus* may act as the origin recognition protein and Mcm helicase recruiter. *Genes Cells*, 2010, 15: 537–552
- 20 Lundgren M, Andersson A, Chen L, et al. Three replication origins in *Sulfolobus* species: synchronous initiation of chromosome replication and asynchronous termination. *Proc Natl Acad Sci USA*, 2004, 101: 7046–7051
- 21 Robinson N P, Blood K A, McCallum S A, et al. Sister chromatid junctions in the hyperthermophilic archaeon *Sulfolobus solfataricus*.

- EMBO J, 2007, 26: 816–824
- 22 Robinson N P, Bell S D. Extrachromosomal element capture and the evolution of multiple replication origins in archaeal chromosomes. *Proc Natl Acad Sci USA*, 2007, 104: 5806–5811
 - 23 Norais C, Hawkins M, Hartman A L, et al. Genetic and physical mapping of DNA replication origins in *Haloferax volcanii*. *PLoS Genet*, 2007, 3:e77
 - 24 Capes M D, Coker J A, Gessler R, et al. The information transfer system of halophilic archaea. *Plasmid*, 2011, 65: 77–101
 - 25 Masai H, You Z, Arai K. Control of DNA replication: regulation and activation of eukaryotic replicative helicase, MCM. *IUBMB Life*, 2005, 57: 323–335
 - 26 Ilves I, Petojevic T, Pesavento J J, et al. Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. *Mol Cell*, 2010, 37: 247–258
 - 27 Chong J P, Hayashi M K, Simon M N, et al. A double-hexamer archaeal minichromosome maintenance protein is an ATP-dependent DNA helicase. *Proc Natl Acad Sci USA*, 2000, 97: 1530–1535
 - 28 Grainge I, Scaife S, Wigley D B. Biochemical analysis of components of the pre-replication complex of *Archaeoglobus fulgidus*. *Nucleic Acids Res*, 2003, 31: 4888–4898
 - 29 Kelman Z, Lee J K, Hurwitz J. The single minichromosome maintenance protein of *Methanobacterium thermoautotrophicum* ΔH contains DNA helicase activity. *Proc Natl Acad Sci USA*, 1999, 96: 14783–14788
 - 30 Shechter D F, Ying C Y, Gautier J. The intrinsic DNA helicase activity of *Methanobacterium thermoautotrophicum* ΔH minichromosome maintenance protein. *J Biol Chem*, 2000, 275: 15049–15059
 - 31 Yoshimochi T, Fujikane R, Kawanami M, et al. The GINS complex from *Pyrococcus furiosus* stimulates the MCM helicase activity. *J Biol Chem*, 2008, 283: 1601–1609
 - 32 Sakakibara N, Kelman L M, Kelman Z. Unwinding the structure and function of the archaeal MCM helicase. *Mol Microbiol*, 2009, 72: 286–296
 - 33 Shin J H, Grabowski B, Kasiviswanathan R, et al. Regulation of minichromosome maintenance helicase activity by Cdc6. *J Biol Chem*, 2003, 278: 38059–38067
 - 34 Kasiviswanathan R, Shin J H, Kelman Z. Interactions between the archaeal Cdc6 and MCM proteins modulate their biochemical properties. *Nucleic Acids Res*, 2005, 33: 4940–4950
 - 35 De Felice M, Esposito L, Pucci B, et al. Biochemical characterization of a CDC6-like protein from the crenarchaeon *Sulfolobus solfataricus*. *J Biol Chem*, 2003, 278: 46424–46431
 - 36 Haugland G T, Shin J H, Birkeland N K, et al. Stimulation of MCM helicase activity by a Cdc6 protein in the archaeon *Thermoplasma acidophilum*. *Nucleic Acids Res*, 2006, 34: 6337–6344
 - 37 Krupovic M, Gribaldo S, Bamford D H, et al. The evolutionary history of archaeal MCM helicases: a case study of vertical evolution combined with hitchhiking of mobile genetic elements. *Mol Biol Evol*, 2010, 27: 2716–2732
 - 38 Fukui T, Atomi H, Kanai T, et al. Complete genome sequence of the hyperthermophilic archaeon *Thermococcus kodakarensis* KOD1 and comparison with *Pyrococcus* genomes. *Genome Res*, 2005, 15: 352–363
 - 39 Sato T, Fukui T, Atomi H, et al. Targeted gene disruption by homologous recombination in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *J Bacteriol*, 2003, 185: 210–220
 - 40 Sato T, Fukui T, Atomi H, et al. Improved and versatile transformation system allowing multiple genetic manipulations of the hyperthermophilic archaeon *Thermococcus kodakaraensis*. *Appl Environ Microbiol*, 2005, 71: 3889–3899
 - 41 Pan M, Santangelo T J, Li Z, et al. *Thermococcus kodakarensis* encodes three MCM homologs but only one is essential. *Nucleic Acids Res*, 2011, 39: 9671–9680
 - 42 Ishino S, Fujino S, Tomita H, et al. Biochemical and genetical analyses of the three mcm genes from the hyperthermophilic archaeon, *Thermococcus kodakarensis*. *Genes Cells*, 2011, 16: 1176–1189
 - 43 Li Z, Santangelo T J, Čuboňová L, et al. Affinity purification of an archaeal DNA replication protein network. *mBio*, 2010, 1: e00221-10
 - 44 Fujikane R, Ishino S, Ishino Y, et al. Genetic analysis of DNA repair in the hyperthermophilic archaeon, *Thermococcus kodakarensis*. *Genes Genet Syst*, 2010, 85: 243–257
 - 45 Xia Q, Hendrickson E L, Zhang Y, et al. Quantitative proteomics of the archaeon *Methanococcus maripaludis* validated by microarray analysis and real time PCR. *Mol Cell Proteomics*, 2006, 5: 868–881
 - 46 Walters A D, Chong J P. An archaeal order with multiple minichromosome maintenance genes. *Microbiology*, 2010, 156: 1405–1414
 - 47 Takayama Y, Kamimura Y, Okawa M, et al. GINS, a novel multi-protein complex required for chromosomal DNA replication in budding yeast. *Genes Dev*, 2003, 17: 1153–1165
 - 48 Makarova K S, Wolf Y I, Mekhedov S L, et al. Ancestral paralogs and pseudoparalogs and their role in the emergence of the eukaryotic cell. *Nucleic Acids Res*, 2005, 33: 4626–4638
 - 49 Marinsek N, Barry E R, Makarova K S, et al. GINS, a central nexus in the archaeal DNA replication fork. *EMBO Rep*, 2006, 7: 539–545
 - 50 Pacek M, Tutter A V, Kubota Y, et al. Localization of MCM2-7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. *Mol Cell*, 2006, 21: 581–587
 - 51 Ogino H, Ishino S, Mayanagi K, et al. The GINS complex from the thermophilic archaeon, *Thermoplasma acidophilum* may function as a homotetramer in DNA replication. *Extremophiles*, 2011, 15: 529–539
 - 52 Oyama T, Ishino S, Fujino S, et al. Architectures of archaeal GINS complexes, essential DNA replication initiation factors. *BMC Biol*, 2011, 9: 28
 - 53 Kamada K, Kubota Y, Arata T, et al. Structure of the human GINS complex and its assembly and functional interface in replication initiation. *Nat Struct Mol Biol*, 2007, 14: 388–396
 - 54 Choi J M, Lim H S, Kim J J, et al. Crystal structure of the human GINS complex. *Genes Dev*, 2007, 21: 1316–1321
 - 55 Chang Y P, Wang G, Bermudez V, et al. Crystal structure of the GINS complex and functional insight into its role in DNA replication. *Proc Natl Acad Sci USA*, 2007, 104: 12685–12690
 - 56 Sanchez-Pulido L, Ponting C P. Cdc45: the missing RecJ ortholog in eukaryotes? *Bioinformatics*, 2011, 27: 1885–1888
 - 57 Li Z, Pan M, Santangelo T J, et al. A novel DNA nuclease is stimulated by association with the GINS complex. *Nucleic Acids Res*, 2011, 39: 6114–6123
 - 58 Krastanova I, Sannino V, Amenitsch H, et al. Structural and functional insights into the DNA replication factor Cdc45 reveal an evolutionary relationship to the DHH family of phosphoesterases. *J Biol Chem*, 2012, 287: 4121–4128
 - 59 Kirk B W, Kuchta R D. Arg304 of human DNA primase is a key contributor to catalysis and NTP binding: primase and the family X polymerases share significant sequence homology. *Biochemistry*, 1999, 38: 7727–7736
 - 60 Desogus G, Onesti S, Brick P, et al. Identification and characterization of a DNA primase from the hyperthermophilic archaeon *Methanococcus jannaschii*. *Nucleic Acids Res*, 1999, 27: 4444–4450
 - 61 Bocquier A, Liu L, Cann I, et al. Archaeal primase: bridging the gap between RNA and DNA polymerase. *Curr Biol*, 2001, 11: 452–456
 - 62 Liu L, Komori K, Ishino S, et al. The archaeal DNA primase: biochemical characterization of the p41-p46 complex from *Pyrococcus furiosus*. *J Biol Chem*, 2001, 276: 45484–45490
 - 63 Lao-Sirieix S H, Bell S D. The heterodimeric primase of the hyperthermophilic archaeon *Sulfolobus solfataricus* possesses DNA and RNA primase, polymerase and 3'-terminal nucleotidyl transferase activities. *J Mol Biol*, 2004, 344: 1251–1263
 - 64 Lao-Sirieix S H, Nookala R K, Roversi P, et al. Structure of the heterodimeric core primase. *Nat Struct Mol Biol*, 2005, 12: 1137–1144

- 65 De Falco M, Fusco A, DeFelice M, et al. The DNA primase is activated by substrates containing a thymine-rich bubble and has a 3'-terminal nucleotidyl-transferase activity. *Nucleic Acids Res*, 2004, 32: 5223–5230
- 66 Ito N, Nureki O, Shirouzu M, et al. Crystal structure of the *Pyrococcus horikoshii* DNA primase-UTP complex: implications for the mechanism of primer synthesis. *Genes Cells*, 2003, 8: 913–923
- 67 Matsui E, Nishio M, Yokoyama H, et al. Distinct domain functions regulating de novo DNA synthesis of thermostable DNA primase from hyperthermophile *Pyrococcus horikoshii*. *Biochemistry*, 2003, 42: 14968–14976
- 68 Ito N, Matsui I, Matsui E. Molecular basis for the subunit assembly of the primase from an archaeon *Pyrococcus horikoshii*. *FEBS J*, 2007, 274: 1340–1351
- 69 Le Breton M, Henneke G, Norais C, et al. The heterodimeric primase from the euryarchaeon *Pyrococcus abyssi*: a multifunctional enzyme for initiation and repair? *J Mol Biol*, 2007, 374: 1172–1185
- 70 Hu J, Guo L, Wu K, et al. Template-dependent polymerization across discontinuous templates by the heterodimeric primase from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *The Nucleic Acids Research* advance online publication 20 December 2011; doi: 10.1093/nar/gkr1256
- 71 Evguenieva-Hackenberg E, Walter P, Hochleitner E, et al. An exosome-like complex in *Sulfolobus solfataricus*. *EMBO Rep*, 2003, 4: 889–893
- 72 Walter P, Klein F, Lorentzen E, et al. Characterization of native and reconstituted exosome complexes from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol Microbiol*, 2006, 62: 1076–1089
- 73 Zuo Z, Rodgers C J, Mikheikin A L, et al. Characterization of a functional DnaG-type primase in archaea: implications for a dual-primase system. *J Mol Biol*, 2010, 397: 664–676
- 74 Wu K, Lai X, Guo X, et al. Interplay between primase and replication factor C in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol Microbiol*, 2007, 63: 826–837
- 75 Murzin A G. OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J*, 1993, 12: 861–867
- 76 Braun K A, Lao Y, He Z, et al. Wold MS Role of protein-protein interactions in the function of replication protein A (RPA): RPA modulates the activity of DNA polymerase alpha by multiple mechanisms. *Biochemistry*, 1997, 36: 8443–8454
- 77 Lin Y L, Chen C, Keshav K F, et al. Dissection of functional domains of the human DNA replication protein complex replication protein A. *J Biol Chem*, 1996, 271: 17190–17198
- 78 Bochkarev A, Bochkareva E, Frappier L, et al. The crystal structure of the complex of replication protein A subunits RPA32 and RPA14 reveals a mechanism for single-stranded DNA binding. *EMBO J*, 1999, 18: 4498–4504
- 79 Bochkareva E, Frappier L, Edwards A M, et al. The RPA32 subunit of human replication protein A contains a single-stranded DNA-binding domain. *J Biol Chem*, 1998, 273: 3932–3936
- 80 Philipova D, Mullen J R, Maniar H S, et al. A hierarchy of SSB protomers in replication protein A. *Genes Dev*, 1996, 10: 2222–2233
- 81 Gomes X V, Wold M S. Structural analysis of human replication protein A. Mapping functional domains of the 70-kD subunit. *J Biol Chem*, 1995, 270: 4534–4543
- 82 Mer G, Bochkarev A, Gupta R, et al. Structural basis for the recognition of DNA repair proteins UNG2, XPA, and RAD52 by replication factor RPA. *Cell*, 2000, 103: 449–456
- 83 Dornreiter I, Erdile L F, Gilbert I U, et al. Interaction of DNA polymerase alpha-primase with cellular replication protein A and SV40 T antigen. *EMBO J*, 1992, 11: 769–776
- 84 Tsurimoto T, Stillman B. Multiple replication factors augment DNA synthesis by the two eukaryotic DNA polymerases, alpha and delta. *EMBO J*, 1989, 8: 3883–3889
- 85 Kenny M K, Lee S H, Hurwitz J. Multiple functions of human single-stranded-DNA binding protein in simian virus 40 DNA replication: single-strand stabilization and stimulation of DNA polymerases alpha and delta. *Proc Natl Acad Sci USA*, 1989, 86: 9757–9761
- 86 Chedin F, Seitz E M, Kowalczykowski S C. Novel homologs of replication protein A in archaea: implications for the evolution of ssDNA-binding proteins. *Trends Biol Sci*, 1998, 23: 273–277
- 87 Kelly T J, Simancek P, Brush G S. Identification and characterization of a single-stranded DNA-binding protein from the archaeon *Methanococcus jannaschii*. *Proc Natl Acad Sci USA*, 1998, 95: 14634–14639
- 88 Kelman Z, Petrokovski S, Hurwitz J. Isolation and characterization of a split B-type DNA polymerase from the archaeon *Methanobacterium thermoautotrophicum* ΔH. *J Biol Chem*, 1999, 274: 28751–28761
- 89 Komori K, Ishino Y. Replication protein A in *Pyrococcus furiosus* is involved in homologous DNA recombination. *J Biol Chem*, 2001, 276: 25654–25660
- 90 Kerr I D, Wadsworth R I, Cubeddu L, et al. Insights into ssDNA recognition by the OB fold from a structural and thermodynamic study of *Sulfolobus* SSB protein. *EMBO J*, 2003, 22: 2561–2570
- 91 Robbins J B, Murphy M C, White B A, et al. Functional analysis of multiple single-stranded DNA-binding proteins from *Methanobacterium acetivorans* and their effects on DNA synthesis by DNA polymerase BI. *J Biol Chem*, 2004, 279: 6315–6326
- 92 Lin Y, Lin L J, Sriratanana P, et al. Engineering of functional replication protein A homologs based on insights into the evolution of oligonucleotide/oligosaccharide-binding folds. *J Bacteriol*, 2008, 190: 5766–5780
- 93 Braithwaite D K, Ito J. Compilation, alignment, and phylogenetic relationships of DNA polymerases. *Nucleic Acids Res*, 1993, 21: 787–802
- 94 Cann I, Ishino Y. Archaeal DNA replication: Identifying the pieces to solve a puzzle. *Genetics*, 1999, 152: 1249–1267
- 95 Ishino Y, Cann I. The euryarchaeotes, a subdomain of archaea, survive on a single DNA polymerase: fact or farce? *Genes Genet Syst*, 1998, 73: 323–336
- 96 Lipps G, Röther S, Hart C, et al. A novel type of replicative enzyme harbouring ATPase, primase and DNA polymerase activity. *EMBO J*, 2003, 22: 2516–2525
- 97 Ohmori H, Friedberg E C, Fuchs R P, et al. The Y-family of DNA polymerases. *Mol Cell*, 2001, 8: 7–8
- 98 Guo C, Kosarek-Stancel J N, Tang T S, et al. Y-family DNA polymerases in mammalian cells. *Cell Mol Life Sci*, 2009, 66: 2363–2381
- 99 Uemori T, Ishino Y, Doi H, et al. The hyperthermophilic archaeon *Pyrodicticum occultum* has two alpha-like DNA polymerases. *J Bacteriol*, 1995, 177: 2164–2177
- 100 Cann I, Ishino S, Nomura N, et al. Two family B DNA polymerases in *Aeropyrum pernix*, an obligate aerobic hyperthermophilic crenarchaeote. *J Bacteriol*, 1999, 181: 5984–5992
- 101 Cann I, Komori K, Toh H, et al. A heterodimeric DNA polymerase: evidence that members of the euryarchaeota possess a novel DNA polymerase. *Proc Natl Acad Sci USA*, 1998, 95: 14250–14255
- 102 Ishino S, Ishino Y. Comprehensive search for DNA polymerase in the hyperthermophilic archaeon, *Pyrococcus furiosus*. *Nucleos Nucleot Nucleic Acids*, 2006, 25: 681–691
- 103 Uemori T, Ishino Y, Toh H, et al. Organization and nucleotide sequence of the DNA polymerase gene from the archaeon *Pyrococcus furiosus*. *Nucleic Acids Res*, 1992, 21: 259–265
- 104 Perler F B, Kumar S, Kong H. Thermostable DNA polymerases. *Adv Protein Chem*, 1996, 48: 377–435
- 105 Imamura K, Fukunaga K, Kawarabayashi Y, et al. Specific interactions of three PCNAs with replication-related proteins in *Aeropyrum pernix*. *Mol Microbiol*, 2007, 64: 308–318
- 106 Ishino Y, Ishino S. Novel DNA polymerases from Euryarchaeota.

- Meth Enzymol, 2001, 334: 249–260
- 107 Gueguen Y, Rolland J L, Lecompte O, et al. Characterization of two DNA polymerases from the hyperthermophilic euryarchaeon *Pyrococcus abyssi*. Eur J Biochem, 2001, 268: 5961–5969
 - 108 Shen Y, Musti K, Hiramoto M, et al. Invariant Asp-1122 and Asp-1124 are essential residues for polymerization catalysis of family D DNA polymerase from *Pyrococcus horikoshii*. J Biol Chem, 2001, 276: 27376–27383
 - 109 Shen Y, Tang X, Matsui I. Subunit interaction and regulation of activity through terminal domains of the family D DNA polymerase from *Pyrococcus horikoshii*. J Biol Chem, 2003, 278: 21247–21257
 - 110 Tang X F, Shen Y, Matsui E, et al. Domain topology of the DNA polymerase D complex from a hyperthermophilic archaeon *Pyrococcus horikoshii*. Biochemistry, 2004, 43: 11818–11827
 - 111 Jokela M, Eskelinen A, Pospiech H, et al. Characterization of the 3' exonuclease subunit DP1 of *Methanococcus jannaschii* replicative DNA polymerase D. Nucleic Acids Res, 2004, 32: 2430–2440
 - 112 Henneke G, Flament D, Hübscher U, et al. The hyperthermophilic euryarchaeota *Pyrococcus abyssi* likely requires the two DNA polymerases D and B for DNA replication. J Mol Biol, 2005, 350: 53–64
 - 113 Castrec B, Laurent S, Henneke G, et al. The glycine-rich motif of *Pyrococcus abyssi* DNA polymerase D is critical for protein stability. J Mol Biol, 2010, 396: 840–848
 - 114 Berquist B R, DasSarma P, DasSarma S. Essential and non-essential DNA replication genes in the model halophilic Archaeon, *Halobacterium* sp. NRC-1. BMC Genet, 2007, 8:31
 - 115 Brochier-Armanet C, Boussau B, Gribaldo S, et al. Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. Nat Rev Microbiol, 2008, 6: 245–252
 - 116 Nunoura T, Takaki Y, Kakuta J, et al. Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. Nucleic Acids Res, 2011, 39: 3204–3223
 - 117 Elkins J G, Podar M, Graham D E, et al. A korarchaeal genome reveals insights into the evolution of the Archaea. Proc Natl Acad Sci USA, 2008, 105: 8102–8107
 - 118 Schleper C, Swanson R V, Mathur E J, et al. Characterization of a DNA polymerase from the uncultivated psychrophilic archaeon *Cenarchaeum symbiosum*. J Bacteriol, 1997, 179: 7803–7811
 - 119 Tahirov T H, Makarova K S, Rogozin I B, et al. Evolution of DNA polymerases: an inactivated polymerase-exonuclease module in Pol epsilon and a chimeric origin of eukaryotic polymerases from two classes of archaeal ancestors. Biol Direct, 2009, 4: 11
 - 120 Connolly B A. Recognition of deaminated bases by archaeal family-B DNA polymerases. Biochem Soc Trans, 2009, 37: 65–68
 - 121 Wardle J, Burgers P M, Cann I K, et al. Uracil recognition by replicative DNA polymerases is limited to the archaea, not occurring with bacteria and eukarya. Nucleic Acids Res, 2008, 36: 705–711
 - 122 Fogg M G, Pearl L H, Connolly B A. Structural basis for uracil recognition by archaeal family B DNA polymerases. Nat Struct Biol, 2002, 9: 922–927
 - 123 Zillig W, Kletzin A, Schleper C, et al. Screening for *Sulfolobales*, their plasmids and their viruses in Icelandic solfatara. Syst Appl Microbiol, 1993, 16: 609–628
 - 124 Lipps G, Röther S, Hart C, et al. A novel type of replicative enzyme harbouring ATPase, primase and DNA polymerase activity. EMBO J, 2003, 22: 2516–2525
 - 125 Lipps G. Molecular biology of the pRN1 plasmid from *Sulfolobus islandicus*. Biochem Soc Trans, 2009, 37: 42–45
 - 126 Lipps G. The replication protein of the *Sulfolobus islandicus* plasmid pRN1. Biochem Soc Trans, 2004, 32: 240–244
 - 127 Soler N, Marguet E, Cortez D, et al. Two novel families of plasmids from hyperthermophilic archaea encoding new families of replication proteins. Nucleic Acids Res, 2010, 38: 5088–5104
 - 128 Moldovan G L, Pfander B, Jentsch S. PCNA, the maestro of the replication fork. Cell, 2007, 129, 665–679
 - 129 Pan M, Kelman L, Kelman Z. The archaeal PCNA proteins. Biochem Soc Trans, 2011, 39: 20–24
 - 130 Cann I, Ishino S, Hayashi I, et al. Functional interactions of a homolog of proliferating cell nuclear antigen with DNA polymerases in Archaea. J Bacteriol, 1999, 181: 6591–6599
 - 131 Ishino Y, Tsurimoto T, Ishino S, et al. Functional interactions of an archaeal sliding clamp with mammalian clamp loader and DNA polymerase δ . Genes Cells, 2001, 6: 699–706
 - 132 Matsumiya S, Ishino S, Ishino Y, et al. Physical interaction between proliferating cell nuclear antigen and replication factor C from *Pyrococcus furiosus*. Genes Cells, 2002, 7: 911–922
 - 133 Matsumiya S, Ishino S, Ishino Y, et al. Intermolecular ion pairs maintain toroidal structure of *Pyrococcus furiosus* PCNA. Prot Sci, 2003, 12: 823–831
 - 134 Tori K, Kimizu M, Ishino S, et al. Both DNA polymerase BI and D from the hyperthermophilic archaeon, *Pyrococcus furiosus* bind to PCNA at the C-terminal PIP box motifs. J Bacteriol, 2007, 189: 5652–5657
 - 135 Cann I, Ishino S, Yuasa M, et al. Biochemical analysis of the replication factor C from the hyperthermophilic archaeon *Pyrococcus furiosus*. J Bacteriol, 2001, 183: 2614–2623
 - 136 Mayanagi K, Miyata T, Oyama T, et al. Three-dimensional electron microscopy of clamp loader small subunit from *Pyrococcus furiosus*. J Struct Biol, 2001, 134: 35–45
 - 137 Ishino S, Oyama T, Yuasa M, et al. Mutational analysis of *Pyrococcus furiosus* replication factor C based on the three-dimensional structure. Extremophiles, 2003, 7: 169–175
 - 138 Oyama T, Ishino Y, Cann I, et al. Atomic structure of the clamp loader small subunit from *Pyrococcus furiosus*. Mol Cell, 2001, 8: 455–463
 - 139 Indiani C, O'Donnell M. The replication clamp-loading machine at work in the three domains of life. Nat Rev Mol Cell Biol, 2006, 7: 751–761
 - 140 Miyata T, Suzuki H, Oyama T, et al. Open clamp structure in the clamp-loading complex visualized by electron microscopic image analysis. Proc Natl Acad Sci USA, 2005, 102: 13795–13800
 - 141 Kelch B A, Makino D L, O'Donnell M, et al. The replication clamp-loading machine at work in the three domains of life. Science, 2011, 334: 1675–1680
 - 142 Warbrick E M. The puzzle of PCNA's many partners. BioEssays, 2000, 22: 997–1006
 - 143 Vivona J B, Kelman Z. The diverse spectrum of sliding clamp interacting proteins. FEBS Lett, 2003, 546: 167–167
 - 144 Nishida H, Mayanagi K, Kiyonari S, et al. Structural determinant for switching between the polymerase and exonuclease modes in the PCNA-replicative DNA polymerase complex. Proc Natl Acad Sci USA, 2009, 106: 20693–20698
 - 145 Shamoo Y, Steitz T A. Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. Cell, 1999, 99: 155–166
 - 146 Franklin M C, Wang J, Steitz T A. Structure of the replicating complex of a pol alpha family DNA polymerase. Cell, 2001, 105: 657–667
 - 147 Mayanagi K, Kiyonari S, Nishida H, et al. Architecture of the DNA polymerase B-proliferating cell nuclear antigen (PCNA)-DNA ternary complex. Proc Natl Acad Sci USA, 2011, 108: 1845–1849
 - 148 Daimon K, Kawarabayashi Y, Kikuchi H, et al. Three proliferating cell nuclear antigen-like proteins found in the hyperthermophilic archaeon *Aeropyrum pernix*: interactions with the two DNA polymerases. J Bacteriol, 2002, 184: 687–694
 - 149 Dionne I, Nookala R K, Jackson S P, et al. A heterotrimeric PCNA in the hyperthermophilic archaeon *Sulfolobus solfataricus*. Mol Cell, 2003, 11: 275–282
 - 150 Imamura K, Fukunaga K, Kawarabayashi Y, et al. Specific interactions of three proliferating cell nuclear antigens with replication-

- related proteins in *Aeropyrum pernix*. Mol Microbiol, 2007, 64: 308–318
- 151 Williams G J, Johnson K, Rudolf J, et al. Structure of the heterotrimeric PCNA from *Sulfolobus solfataricus*. Acta Crystallogr Sect F Struct Biol Cryst Commun, 2006, 62: 944–948
 - 152 Doré A S, Kilkenny M L, Jones S A, et al. Structure of an archaeal PCNA1-PCNA2-FEN1 complex: elucidating PCNA subunit and client enzyme specificity. Nucleic Acids Res, 2006, 34: 4515–4526
 - 153 Ladner J E, Pan M, Hurwitz J, et al. Crystal structures of two active proliferating cell nuclear antigens (PCNAs) encoded by *Thermococcus kodakaraensis*. Proc Natl Acad Sci USA, 2011, 108: 2711–2716
 - 154 Chen Y H, Kocherginskaya S A, Lin Y, et al. Biochemical and mutational analyses of a unique clamp loader complex in the archaeon *Methanosarcina acetivorans*. J Biol Chem, 2005, 280: 41852–41863
 - 155 Chen Y H, Lin Y, Yoshinaga A, et al. Molecular analyses of a three-subunit euryarchaeal clamp loader complex from *Methanosarcina acetivorans*. J Bacteriol, 2009, 191: 6539–6549
 - 156 Chia N, Cann I, Olsen G J. Evolution of DNA replication protein complexes in eukaryotes and Archaea. PLoS ONE, 2010, 5: e10866
 - 157 Wilkinson A, Day J, Bowater R. Bacterial DNA ligases. Mol Microbiol, 2001, 40: 1241–1248
 - 158 Tomkinson A E, Vijayakumar S, Pascal J M, et al. DNA ligases: structure, reaction mechanism, and function. Chem Rev, 2006, 106: 687–699
 - 159 Poidevin L, MacNeill S A. Biochemical characterisation of LigN, an NAD⁺-dependent DNA ligase from the halophilic euryarchaeon *Haloflex volcanii* that displays maximal *in vitro* activity at high salt concentrations. BMC Mol Biol, 2006, 7: 44
 - 160 Benarroch D, Shuman S. Characterization of mimivirus NAD⁺-dependent DNA ligase. Virology, 2006, 353: 133–143
 - 161 Kletzin A. Molecular characterisation of a DNA ligase gene of the extremely thermophilic archaeon *Desulfurolobus ambivalens* shows close phylogenetic relationship to eukaryotic ligases. Nucleic Acids Res, 1992, 20: 5389–5396
 - 162 Nakatani M, Ezaki S, Atomi H, et al. A DNA ligase from a hyperthermophilic archaeon with unique cofactor specificity. J Bacteriol, 2000, 182: 6424–6433
 - 163 Rolland J L, Gueguen Y, Persillon C, et al. Characterization of a thermophilic DNA ligase from the archaeon *Thermococcus fumicolans*. FEMS Microbiol Lett, 2004, 236: 267–273
 - 164 Kim Y J, Lee H S, Bae S S., et al. Cloning, expression, and characterization of a DNA ligase from a hyperthermophilic archaeon *Thermococcus* sp. Biotechnol Lett, 2006, 28: 401–407
 - 165 Ferrer M, Golyshina O V, Belouqui A, et al. A purple acidophilic di-feric DNA ligase from *Ferroplasma*. Proc Natl Acad Sci USA, 2008, 105: 8878–8883
 - 166 Jeon S J, Ishikawa K. A novel ADP-dependent DNA ligase from *Aeropyrum pernix* K1. FEBS Lett, 2003, 550: 69–73
 - 167 Seo M S, Kim Y J, Choi J J, et al. Cloning and expression of a DNA ligase from the hyperthermophilic archaeon *Staphylothermus marinus* and properties of the enzyme. J Biotechnol, 2007, 128: 519–530
 - 168 Sun Y, Seo M S, Kim J H, et al. Novel DNA ligase with broad nucleotide cofactor specificity from the hyperthermophilic crenarchaeon *Sulfolobococcus zilligii*: influence of ancestral DNA ligase on cofactor utilization. Environ Microbiol, 2008, 10: 3212–3224
 - 169 Keppetipola N, Shuman S. Characterization of a thermophilic ATP-dependent DNA ligase from the euryarchaeon *Pyrococcus horikoshii*. J Bacteriol, 2005, 187: 6902–6908
 - 170 Kiyonari S, Takayama K, Nishida H, et al. Identification of a novel binding motif in *Pyrococcus furiosus* DNA ligase for the functional interaction with proliferating cell nuclear antigen. J Biol Chem, 2006, 281: 28023–28032
 - 171 Nishida H, Kiyonari S, Ishino Y, et al. The closed structure of an archaeal DNA ligase from *Pyrococcus furiosus*. J Mol Biol, 2006, 360: 956–967
 - 172 Levin D S, Bai W, Yao N, et al. An interaction between DNA ligase I and proliferating cell nuclear antigen: implications for Okazaki fragment synthesis and joining. Proc Natl Acad Sci USA, 1997, 94: 12863–12868
 - 173 Jónsson Z O, Hindges R, Hübscher U. Regulation of DNA replication and repair proteins through interaction with the front side of proliferating cell nuclear antigen. EMBO J, 1998, 17: 2412–2425
 - 174 Tom S, Henricksen L A, Park M S, et al. DNA ligase I and proliferating cell nuclear antigen form a functional complex. J Biol Chem, 2001, 276: 24817–24825
 - 175 Kiyonari S, Kamigochi T, Ishino Y. A single amino acid substitution in the DNA-binding domain of *Aeropyrum pernix* DNA ligase impairs its interaction with proliferating cell nuclear antigen. Extremophiles, 2007, 11: 675–684
 - 176 Kiyonari S, Uchimura M, Shirai T, et al. Physical and functional interactions between uracil-DNA glycosylase and proliferating cell nuclear antigen from the euryarchaeon *Pyrococcus furiosus*. J Biol Chem, 2008, 283: 24185–24193
 - 177 Pascal J M, O'Brien P J, Tomkinson A E, et al. Human DNA ligase I completely encircles and partially unwinds nicked DNA. Nature, 2004, 432: 473–478
 - 178 Pascal J M, Tsodikov O V, Hura G L, et al. A flexible interface between DNA ligase and PCNA supports conformational switching and efficient ligation of DNA. Mol Cell, 2006, 24: 279–291
 - 179 Mayanagi K, Kiyonari S, Saito M, et al. Mechanism of replication machinery assembly as revealed by the DNA ligase-PCNA-DNA complex architecture. Proc Natl Acad Sci USA, 2009, 106: 4657–4662
 - 180 Liu Y, Kao H I, Bambara R A. Flap endonuclease 1: a central component of DNA metabolism. Annu Rev Biochem, 2004, 73: 589–615
 - 181 Zheng L, Jia J, Finger L D, et al. Functional regulation of FEN1 nuclease and its link to cancer. Nucleic Acids Res, 2011, 39: 781–794
 - 182 Rossi M L, Bambara R A. Reconstituted Okazaki fragment processing indicates two pathways of primer removal. J Biol Chem, 2006, 281: 26051–26061
 - 183 Tom S, Henricksen L A, Bambara R A. Mechanism whereby proliferating cell nuclear antigen stimulates flap endonuclease 1. J Biol Chem, 2000, 275: 10498–10505
 - 184 Sakurai S, Kitano K, Yamaguchi H, et al. Structural basis for recruitment of human flap endonuclease 1 to PCNA. EMBO J, 2005, 24: 683–693
 - 185 Shen B, Qiu J, Hosfield D, et al. Flap endonuclease homologs in archaeobacteria exist as independent proteins. Trends Biochem Sci, 1998, 23: 171–173
 - 186 Hwang K Y, Baek K, Kim H Y, et al. The crystal structure of flap endonuclease-1 from *Methanococcus jannaschii*. Nat Struct Biol, 1998, 5: 707–713
 - 187 Hosfield D J, Mol C D, Shen B, et al. Structure of the DNA repair and replication endonuclease and exonuclease FEN-1: coupling DNA and PCNA binding to FEN-1 activity. Cell, 1998, 95: 135–146
 - 188 Matsui E, Musti K V, Abe J, et al. Molecular structure and novel DNA binding sites located in loops of flap endonuclease-1 from *Pyrococcus horikoshii*. J Biol Chem, 2002, 277: 37840–37847
 - 189 Chapados B R, Hosfield D J, Han S, et al. Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. Cell, 2004, 116: 39–50
 - 190 Matsui E, Kawasaki S, Ishida H, et al. Thermostable flap endonuclease from the archaeon, *Pyrococcus horikoshii*, cleaves the replication fork-like structure endo/exonucleolytically. J Biol Chem, 1999, 274: 18297–18309
 - 191 Matsui E, Abe J, Yokoyama H, et al. Aromatic residues located close to the active center are essential for the catalytic reaction of flap endonuclease-1 from hyperthermophilic archaeon *Pyrococcus horikoshii*. J Biol Chem, 2004, 279: 16687–16696
 - 192 Beattie T R, Bell S D. Molecular machines in archaeal DNA replication. Curr Opin Chem Biol, 2011, 15: 614–619

- 193 Wang Y, Prosen D E, Mei L, *et al.* A novel strategy to engineer DNA polymerases for enhanced processivity and improved performance *in vitro*. *Nucleic Acids Res*, 2004, 32: 1197–1207
- 194 Kawamura A, Ishino Y, Ishino S. Biophysical analysis of PCNA from *Pyrococcus furiosus*. *J Jap Soc Extremophiles*, 2012, in press
- 195 Ishino S, Kawamura A, Ishino Y. Application of PCNA to processive PCR by reducing the stability of its ring structure. *J Jap Soc Extremophiles*, 2012, in press
- 196 Shendure J, Ji H. Next-generation DNA sequencing. *Nat Biotechnol*, 2008, 26: 1135–1145
- 197 Ansorge W J. Next-generation DNA sequencing techniques. *N Biotechnol*, 2009, 25: 195–203
- 198 Metzker M L. Sequencing technologies - the next generation. *Nat Rev Genet*, 2010, 11: 31–46
- 199 Jackson B R, Noble C, Lavesa-Curto M, *et al.* Characterization of an ATP-dependent DNA ligase from the acidophilic archaeon “*Ferroplasma acidarmanus*” Fer1. *Extremophiles*, 2007, 11: 315–327
- 200 Sriskanda V, Kelman Z, Hurwitz J, *et al.* Characterization of an ATP-dependent DNA ligase from the thermophilic archaeon *Methanobacterium thermoautotrophicum*. *Nucleic Acids Res*, 2000, 28: 2221–2228
- 201 Lai X, Shao H, Hao F, *et al.* Biochemical characterization of an ATP-dependent DNA ligase from the hyperthermophilic crenarchaeon *Sulfolobus shibatae*. *Extremophiles*, 2002, 6: 469–477
- 202 Bezudnova E Y, Kovalchuk M V, Mardanov A V, *et al.* Overexpression, purification and crystallization of a thermostable DNA ligase from the archaeon *Thermococcus* sp. 1519. *Acta Crystallogr Sect F Struct Biol Cryst Commun*, 2009, 65: 368–371

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.